

us-09-486-167a-l.rge

Mon Oct 22 14:46:51 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: October 22, 2001, 10:42:40 ; Search time 1293.29 Seconds
(without alignments)
9627.800 Million cell updates/sec

Title: US-09-486-167A-1
Perfect score: 805
Sequence: 1 gccaggagcgagtggaag.....ttgtgtttgcggaaaaaaa 805

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues 2688314
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_ov.*
8: gb_pat1.*
9: gb_pat2.*
10: gb_ph.*
11: gb_ph.*
12: gb_ph.*
13: gb_ph.*
14: gb_ph.*
15: gb_ph.*
16: gb_ph.*
17: gb_ph.*
18: gb_ph.*
19: gb_ph.*
20: gb_ph.*
21: gb_ph.*
22: gb_ph.*
23: gb_ph.*
24: gb_ph.*
25: gb_ph.*
26: gb_ph.*
27: gb_ph.*
28: gb_ph.*
29: gb_ph.*
30: gb_ph.*
31: gb_ph.*
32: gb_ph.*
33: gb_ph.*
34: gb_ph.*
35: gb_ph.*
36: gb_ph.*
37: gb_ph.*
38: gb_ph.*
39: gb_ph.*
40: gb_ph.*
41: gb_ph.*
42: gb_ph.*
43: gb_ph.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vi.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vil.*
59: gb_vil.*
60: gb_vil.*
61: gb_vil.*
62: gb_vil.*
63: gb_vil.*
64: gb_vil.*
65: gb_vil.*
66: gb_vil.*
67: gb_vil.*
68: gb_vil.*
69: gb_vil.*
70: gb_vil.*
71: gb_vil.*
72: gb_vil.*
73: gb_vil.*
74: gb_vil.*
75: gb_vil.*
76: gb_vil.*
77: gb_vil.*
78: gb_vil.*
79: gb_vil.*
80: gb_vil.*
81: gb_vil.*
82: gb_vil.*
83: gb_vil.*
84: gb_vil.*
85: gb_vil.*
86: gb_vil.*
87: gb_vil.*
88: gb_vil.*
89: gb_vil.*
90: gb_vil.*
91: gb_vil.*
92: gb_vil.*
93: gb_vil.*
94: gb_vil.*
95: gb_vil.*
96: gb_vil.*
97: gb_vil.*
98: gb_vil.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	805	100.0	805	9	A99097	A99097 Sequence 1
2	805	100.0	805	88	AF110731	AF110731 Homo sapi
3	800.2	99.4	840	89	AF124993	AF124993 Homo sapi
4	788.8	98.0	875	89	AF242525	AF242525 Homo sapi
5	778.4	96.7	788	89	AF231705	AF231705 Homo sapi
6	759.8	94.4	835	89	AF112212	AF112212 Homo sapi
7	690.2	85.7	780	89	AF197952	AF197952 Homo sapi
8	650	80.7	725	88	AF110736	AF110736 Cercopit

/product="peroxisomal membrane protein 20"
 /protein_id="AAF27531.1"
 /db_xref="GI:6746355"
 /translation="MAPIKVGDAIPAVEFEGEPGNKVNLAELFKGKKGVLFVGPQAF
 TPQCSYTHLPFGEQAEALKAQVQVACLSDNDAFTVGEWRAHKAEGKVRLLADPT
 GAFKGTDLDDSLVSIIFGNRLKRFNMVQDGVKALNVEPDGTGLTCLAPNIIS
 OL"
 BASE COUNT 166 a 214 c 289 g 171 t
 ORIGIN

Query Match 99.4%; Score 800.2; DB 89; Length 840;
 Best Local Similarity 99.6%; Pred. No. 3e-192;
 Matches 802; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 gccagagcgagtggaagtggcgtggcggtatggactagctggcgtgtgcgc 60
 Db 34 GCCAGAGCGGAGTGGAACTGGCGTGGCGGATGGGACTAGCTGGCGTGGCGC 93
 QY 61 ctgagacgtcagcggtctatactctcgttggcggtcagctgcgcgcagc 120
 Db 94 CTGAGACGCTCAGCGGCTATATCTCGTGGTGGCGCGCTCAGTCTGGCGAGCG 153
 QY 121 gcagcaagcagctgcagtgagggagtgggcgctctggcggttcagctgcgcgcagc 180
 Db 154 GCAGCAAGACGCTGAGTGAAGGAGTGGCGCTCTGGCGGGTCCGCACTTCACAGCA 213
 QY 181 gccgctgcagccatggcccaatacgaagtggagtgccatcccgagctgagtggtt 240
 Db 214 GCCGCTGCAGCCATGGCCCAATCAGTGGGAGTGGCATCCAGCAGTGGAGTGT 273
 QY 241 aaaggagcagggaaacaaaggtgaacctggcagagcttcaagggaagaaggtgtg 300
 Db 274 GAAGGGAGCGAGGAAACAAAGGTGAACCTGGCAGAGCTTTCAAGGGCAAGAGGTGT 333
 QY 301 ctgtttgagttctctgggcttcacccctgagtgatgttcaagacacacactgccaggttt 360
 Db 334 CTGTTTGGAGTCTCTGGGCTTCACCCCTGGATGTTTCAAGACACACACTGCCAGGTTT 393
 QY 361 gtgagcagctgagctctgaagcgaaggttccaggttccaggttccaggttccaggtt 420
 Db 394 GTGAGCAGCTGAGGCTCTGAAGCAGGAGTCCAGGTGGTGGCTCTGTGAGTGT 453
 QY 421 aatgtgctcttctgactggcagtgaggcagcagcagcagcagcagcagcagcagcagc 480
 Db 454 AATGATGCTTGTGACTGGCAGTGGGCGGAGCCACACAGGCGAAGCAAGTTCG 513
 QY 481 ctctgctgatccactggggtcttgggaagagagacagacttactagatgattcg 540
 Db 514 CTCTGGCTGATCCCACTGGGCGCTTTGGGAAGGAGAGACACTATTACTAGATGATT 573
 QY 541 ctggtgtccatcttgggaatcgactctcaagaggttccatgtgttacagagtggtc 600
 Db 574 CTGGTGTCCATCTTGGGAATCGAGCTCTCAAGAGGTCTTCCATGTGTGTACAGGATG 633
 QY 601 atagtgaagccctgaatgtggaaccagatggaacaggttccatgtgttacagagtggtc 660
 Db 634 ATAGTGAAGCCCTGAATGTGGAACAGATGGACAGGCTCACCTGAGCCCTGGCACCC 693
 QY 661 aatatcatcagagctctgagggcctgagggcagattacttctccacccctccctatct 720
 Db 694 AATATCTCTACAGCTCTGAGGCCCTGGGCCAGATTAATCTCTCCACCCCTCCCTATCT 753
 QY 721 cactgtccagcctgtgtggggtcctgcaattggaatgttggccagattcttctccacccctcc 780
 Db 754 CACCTGCCAGCGTGTGTGGGGCCCTGCAATTGGAATGTTGGCCAGATTTCTGCAATA 813
 QY 781 aacactgtgtgttgcggaaaaaa 805
 Db 814 AACACTGTGTGTGGGGCCCAAAA 838

RESULT 4

AF242525
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AF242525 875 bp mRNA
 Homo sapiens hypothetical protein SBB110 mRNA, complete cds.
 AF242525.1 GI:9802047
 human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 875)

Zhang, W., Li, N., Wan, T. and Cao, X.

Direct Submission

Submitted (08-MAR-2000) Department of Immunology, Second Military

Medical University & Shanghai Brilliance Biotechnology Institute,

800 Xiangyin Rd., Shanghai 200433, P.R. China

Location/Qualifiers

1..875

/organism="Homo sapiens"

/db_xref="taxon:9606"

36..680

/codon_start=1

/product="hypothetical protein SBB110"

/protein_id="AAF99605.1"

/db_xref="GI:9802048"

/translation="MGLAGYCALRRSAGYILVGGAGGQSAARRCSEGEWASGVFR

SFSRAAAMAPIKVGDAIPAVEFEGEPGNKVNLAELFKGKKGVLFVGPQAFPGCSK

THEPCFVEQAEALKAQVQVACLSDNDAFTVGEWRAHKAEGKVRLLADPTGAFGKE

TDLLDDSLVSIIFGNRLKRFNMVQDGVKALNVEPDGTGLTCLAPNIISQL"

BASE COUNT 230 a 206 c 272 g 167 t

ORIGIN

Query Match 98.0%; Score 788.8; DB 89; Length 875;
 Best Local Similarity 99.7%; Pred. No. 2.3e-189;
 Matches 790; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 14 gtgaaagtggcgtggcggttatggactagctggcgtgtgcgcctgagagctcag 73
 Db 13 GTGAAAGTGGCGGTGGCGGGTATGGGACTAGCTGGCGTGTGCCCTGTGAGCGCTCAG 72
 QY 74 cgggtctatactctcgttggcggtcgtcagctctcggcagcagcagcagcagcagc 133
 Db 73 CGGGCTATATACCTCGTGGTGGCGCGGCTCAGCTCTGCGGCAGCGCAGCAGCGT 132
 QY 134 gcagtgaagagaagtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 193
 Db 133 GCAGTGAAGAGAGTGGCGGTCTGCGGGGTCTGCGGGGTCTGCGGGGTCTGCGGGGT 192
 QY 194 tggccccaatcaagtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 253
 Db 193 TGCGCCCAATCAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 252
 QY 254 ggaacaagtgaaacctggcagagctgttcaaggggcaagggggtgtgtgtgtgtgtgt 313
 Db 253 GGAACAAGTGAACCTGGCAGAGTGTTCAGGGCAAGAAGGGTGTGTGTGTGTGTGT 312
 QY 314 ctggggcttccacccctggatgttcccaagacacacacacacacacacacacacacac 373
 Db 313 CTGGGGCTTCCACCCCTGGATGTTCCCAAGACACACACTGCCAGGGTGTGTGAGCAG 372
 QY 374 aggtcttgaagcccaagggagtgccagtggtggtggtggtggtggtggtggtggtggt 433
 Db 373 AGGCTCTGAAGCCCAAGGGAGTCCAGTGTGCTGTCTGAGTGTCTGAGTGTCTGAGT 432
 QY 434 tgactgtcagtggtggcgagcccaagcggaagcggaagcggaagcggaagcggaagcg 493
 Db 433 TGACTGCGAGTGGCGAGCCAGCCACCAAGCGGAGGCAAGGCTCGGCTCGGCTGATC 492
 QY 494 ccactgtggccttggaaagagagacagacttattactagatgattcgtggtgtccatct 553
 Db 493 CCACTGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCCTGCTGTCCATCT 552

QY 554 ttgggaatcagactctcaagagttctccatgtgtgtacagagtgatgacatgtaagggccc 613
 Db 553 TTGGGAATCAGAGCTCTCAAGAGGTTCTCCATGTGTGTACAGGATGCATAGTGAAGGCC 612

QY 614 tgaatgtggaacagatggaacagagctcaactgagcctggcaccacatatactatc 673
 Db 613 TGAATGTGGAACAGATGGACAGGCTCACTCCAGCCTGGCACCACCAATATCATCTCAC 672

QY 674 agctctgagcctgggagagattactctccacccctctctctcactctccagcc 733
 Db 673 AGCTCTGAGCCTGGGCCAGATTACTTCTCCACCCCTCCCTATCTCACCTGCCAGCC 732

QY 734 ctgtgtgtgggcccctgcaattggaatgtgtggccagattttctgcaataaacacttgtgtt 793
 Db 733 CTGTGTGTGGGCCCTGCAATTGGAATGTTGGCCAGATTTCTGCAATAAACACTTGTGGTT 792

QY 794 tgcggaaaaaaa 805
 Db 793 TGGCGCCCAAAAA 804

RESULT 5
 AF231705 AF231705 788 bp mRNA PRI 27-JUN-2000
 LOCUS Homo sapiens Alu co-repressor 1 (ACR1) mRNA, complete cds.
 DEFINITION AF231705
 ACCESSION AF231705
 VERSION AF231705.1 GI:8745393
 KEYWORDS
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 788)
 Kropotov, A., Sedova, V., Ivanov, V., Sazeeva, N., Tomilin, A.,
 Krut'ina, R., Oei, S.L., Griesenbeck, J., Buchlow, G. and Tomilin, N.
 A novel human DNA-binding protein with sequence similarity to a
 subfamily of redox proteins which is able to repress
 RNA-polymerase-III-driven transcription of the Alu-family
 retroposons in vitro

AUTHORS Kropotov, A., Sedova, V., Ivanov, V., Sazeeva, N., Tomilin, A.,
 Krut'ina, R., Oei, S.L., Griesenbeck, J., Buchlow, G. and Tomilin, N.

TITLE A novel human DNA-binding protein with sequence similarity to a
 subfamily of redox proteins which is able to repress
 RNA-polymerase-III-driven transcription of the Alu-family
 retroposons in vitro

JOURNAL Eur. J. Biochem. 260 (2), 336-346 (1999)
 MEDLINE 99195471
 PUBMED 10095767

REFERENCE 2 (bases 1 to 788)
 Kropotov, A.V. and Tomilin, N.V.
 Direct Submission
 Submitted (04-FEB-2000) Chromosome Stability, Institute of Cytology
 RAS, Tikhonovskii Av 4, St.Petersburg 194064, Russia

FEATURES
 source location/Qualifiers
 1..788
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q13"
 1..788
 /gene="ACR1"
 20..664
 /gene="ACR1"
 /codon_start=1
 /product="Alu co-repressor 1"
 /protein_id="AAF78899.1"
 /db_xref="GI:8745394"
 /translation="MGLAGVCAILRRSAGYILVGGAGGQSAARRCSEGNWAGGVR
 SFRRAAAMAPIKVDAIPAVEFEGEPGNKNLAELFKKGVLFGVPGATPGCSK
 THLPGFVEQAEALAKGVQVAVCLSVNDAFTVGWRAHKAEGKVRLLADPTGAFGE
 TDLLDLSLVSIFGNRLKRFMSMVODGIVKALNVEPDGTGLTCSLAPNIISOL"

BASE COUNT 153 a 201 c 268 g 166 t
 ORIGIN

Query Match 96.7%; Score 778.4; DB 89; Length 788;
 Best Local Similarity 99.9%; Pred. No. 1e-186;
 Matches 779; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 agtggccctggggcggtatgggactagctggcgtgtgctgcccctgagcgcgtcagcgggc 78
 Db 2 AGTGGCCCTGGGGCGGTTATGGGACTAGCTGGCGTGTGGCCCTGAGACGCTCAGCGGCG 61

QY 79 tatataactcgtcggtggggcggtcagctcgccgagcagcgagcagcagcagcagcagc 138
 Db 62 TATATACTCGTGGGTGGGGCGGCGGTGAGTCTGGCGAGCGGCGAGCAAGAGAGGTTGCA 121

QY 139 gaagagagagtggtcgtcgtggcggtggtccgcaggtttccagcagagcgcgtcgagccatg 198
 Db 122 GAAGGAGAGTGGGCTGTGGGGGGGTCCGCAGTTCACGAGAGCGCTGCAGCCATGCC 181

QY 199 ccaatacgaagtggtggagatgcaatccacagcagtgaggggtgtttgaaggggagccaggg 258
 Db 182 CCAATCAAGGTGGGAGATGCCATCCACAGTGGAGGTGTTTGAAGGGGAGCCAGGGAAC 241

QY 259 aaggtgaacctggcagagctgttcaaggcgcaagaggggtgtgtctgttggagttcctgg 318
 Db 242 AAGGTGAACCTGGCAGAGCTGTTCAAGGCAAGAGGGTGTGCTGTTTGGAGTTCTCTGG 301

QY 319 gcttcacccctggatgttccaagacacacacctggcagggtttgtggagcaggtgaggt 378
 Db 302 GCCTTACCCCTGGATGTTCCAAGACACACCTGCCAGGGTTTGTGAGCAGGCTCAGGCT 361

QY 379 ctgaagggcacaagggagtcaggtggtgctgtctgagtggttaagtgaatgagcctttgact 438
 Db 362 CTGAAGGCCAAGGAGTCCAGGTGGCTGCTCTGAGTGTAAATGATGCCCTTTGTGACT 421

QY 439 ggcagatggggcgagccacacagcaggaaggaaggttcgctcctgagctgagctccact 498
 Db 422 GCGAGTGGGGCGGAGCCACCAAGGCGGAAGGTTTGGCTCCTGCTGCTGATGCCACT 481

QY 499 ggggctttgggaaggagacagactattactagatgattcgtggtggtccactctttgg 558
 Db 482 GGGGCTTTGGGAAGGAGACACACTATTACTAGATGATTCCTGCTGCTGCTCCTTTGG 541

QY 559 aatcagctctcaagaggttccatggtggtggtggtggtggtggtggtggtggtggtggt 618
 Db 542 AATCGACGCTCTCAAGAGGTCTCCATGTTGTGTACAGGATGGCATAGTAGAAGCCCTGA 601

QY 619 gfggaaccagatggcacagcctcacctcagcctgagcccaacacacacacacacacac 678
 Db 602 GTGGAACACAGATGGCACAGGCTCAGCTGCGAGCCCTGGCAACCAATATCATCTCAG 661

QY 679 tgaggccctgggcccagattacttctccacccctccctatctcactgcccagccctg 738
 Db 662 TGAGGCCCTGGCGACAGATTACTTCTCCACCCTCCCTATCTCAGCTGCCAGCCCTGTG 721

QY 739 ctggggccctgcaattggaatgttgccagattctgcaataaacactgtgtgtggtggt 798
 Db 722 CTGGGCCCTGCAATTGGAATGTTGGCCAGATTTCTGCAATAAACACTTGTGTTGGCG 781

RESULT 6
 AF112212 AF112212 835 bp mRNA PRI 13-DEC-1999
 LOCUS Homo sapiens putative peroxisomal antioxidant enzyme mRNA, complete
 DEFINITION cds.
 ACCESSION AF112212
 VERSION AF112212.1 GI:6563211
 KEYWORDS
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 835)
 Ren, S., Shi, J., Huang, C., Jiang, C., Li, Y., Zhou, J., Yu, Y., Xu, S.,
 Wang, Y., Fu, G., Chen, Z. and Han, Z.

AUTHORS A novel gene expressed in human adrenal gland

TITLE Unpublished

JOURNAL

REFERENCE 2 (bases 1 to 835)
 Ren, S., Shi, J., Huang, C., Jiang, C., Li, Y., Zhou, J., Yu, Y., Xu, S.,
 Wang, Y., Fu, G., Chen, Z. and Han, Z.


```

Db 254 GCCGCTGAGCCATGCGCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTTT 313
Qy 241 gaagggagcaggaagaagtgaaactgagcagagctgttcaaggaggaagaaggtgtg 300
Db 314 GAAGGGAGCAGGAGACAGGTGAACCTGGCAGAGCTGTTCAAGGGCAGAGAGGTGTG 373
Qy 301 ctgtttggagttctctggggccttcaaccctggatgtttccaaagacaacacctgccagggtt 360
Db 374 CTGTTTGGAGTTCCTGGGGCCTTCAACCCCTGGATGTTTCCAAGACACACACCTGCCAGGTTT 433
Qy 361 gtgagcagctaggtctgaagccaaggaagtcaggtgtgctctcagttgtt 420
Db 434 GTGAGCAGGCTGAGGCTCTGAAGCCCAAGGAGTGCAGGTGGTGGCTCTCTGAGTGT 493
Qy 421 aatgatgcttctgactggcagatggggcggagccacacagcggaagcaaggttcgg 480
Db 494 AATGATGCTTGTGACTGGGAGCTGGGGCCGAGCCACAAAGGGGGAAGGCAAGGTTCGG 553
Qy 481 ctctggtctatccactggggccttgggaaggagacagacttattactagatctg 540
Db 554 CTCCTGGCTGATCCACTGGGGCCTTGGGAAGGAGACAGACTTATTACTAGATGATCG 613
Qy 541 ctggtccactcttggaaatcagctctcaagaggttctccatggtgtacagatggc 600
Db 614 CTGTTGTCCTCACTTTGGGAATCGAGCTCTCAAGAGGTCTCCATGTTGGTACAGGATGC 673
Qy 601 atagtgaaggccctgaatgtgaaccagatggcacaggtcctcacctcagcctggcaccc 660
Db 674 ATAGTGAAGCCCTGAATGTGAACCATGATGGCACAGGCTCAGCTCAGGCTGGCACCC 733
Qy 661 aatatcatctcaagctctgagggccctggccagattactctccacc 709
Db 734 AATATCATCTCACAGCTCTGA--CCCTGGGCCAGATTACTTCTCTCCACC 780

RESULT 8
AF110736 725 bp mRNA PRI 27-SEP-2000
LOCUS Cercopithecus aethiops peroxiredoxin 5 (PRDX5) mRNA, complete cds.
DEFINITION
ACCESSION AF110736
VERSION AF110736.2 GI:10305337
KEYWORDS African green monkey.
SOURCE Cercopithecus aethiops
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopithecinae; Cercopithecus.
REFERENCE
1 (bases 1 to 725)
Knoops,B., Clippe,A., Bogard,C., Arsalane,K., Wattiez,R.,
Hermans,C., Duconsellie,E., Falmagne,P. and Bernard,A.
Cloning and characterization of AOEB166, a novel mammalian
antioxidant enzyme of the peroxiredoxin family
J. Biol. Chem. 274 (43), 30451-30458 (1999)
99452929
PUBMED 10521424
REFERENCE
2 (bases 1 to 725)
Knoops,B. and Cherif,H.
Cloning and characterization of COS-7 AOEB166/PRDX5
Unpublished
3 (bases 1 to 725)
Knoops,B., Wattiez,R., Falmagne,P., Hermans,C. and Bernard,A.
Direct Submission
Submitted (04-DEC-1998) Department of Biology, Catholic University
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
4 (bases 1 to 725)
Knoops,B., Wattiez,R., Falmagne,P., Hermans,C. and Bernard,A.
Direct Submission
Submitted (26-SEP-2000) Department of Biology, Catholic University
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
REMARK
COMMENT Sequence update by submitter
On Sep 26, 2000 this sequence version replaced gi:10129963.
FEATURES
Location/Qualifiers
1..725
source

```

```

/organism="Cercopithecus aethiops"
/db_xref="taxon:9534"
/cell_line="COS-7"
1..725
/gene="PRDX5"
5..552
/gene="PRDX5"
/note="AOEB166; PRDX5; alkyl hydroperoxide reductase;
thioredoxin peroxidase; mitochondrial, peroxisomal, and
cytosolic"
/codon_start=1
/product="peroxiredoxin 5"
/protein_id="AAG13453.2"
/db_xref="GI:10305338"
/translation="MGLAGVGLRRSAGYILGGAARSAVATAAARRRSEGGWASGGV
RSFSRAAAMAPIKVGDAPAVEVEFEGEPGNKVNLAELFKGKGVLFQVGPATPGGS
KTHLPFGVEQAEALKAQGVQLACLNVNDAFTWEMGRAHKAEGKRLVLLADPTGAFGK
ETDLLDLSLVSIIFGNRLKRFMSVWQDGIYKALNVPDGTGLTCLAPSIISQL"
164
/misc_feature
/gene="PRDX5"
/note="probable alternative translation start site"
BASE COUNT 141 a 189 c 245 g 150 t
ORIGIN
Query Match 80.7%; Score 650; DB 88; Length 725;
Best Local Similarity 95.3%; Pred. No. 3.3e-154;
Matches 694; Conservative 0; Mismatches 25; Indels 9; Gaps 2;
QY 33 gggatggactagctggcgtgtgcccctgagacgtcagcgactatatactcgtcgg 92
|||||
Db 1 GGGTATGGGACTAGCTGGCGTGTGCTCTGAGACGCTCAGCGGGCTATATAC---TCGG 57
|||||
QY 93 tggggccggcggcgtcagctctgcggcagc-----gcagcaagacgctgagtgaggaga 146
|||||
Db 58 TGGGGCCGCCCTCAGTCTGTGGCAGCAGCAGCAGCAGCAGCGAAGTGAAGGAGG 117
|||||
QY 147 gtggcgctgtgcgggggtccgcagtttcagcagagcgcctgcagccatggcccaatcaa 206
|||||
Db 118 GTGGCGCTGTGGCGGGTCCGAGTTTCAGCAGAGCGCTGCAGCCATGGCCCGATCAA 177
|||||
QY 207 ggtggagatgccatccacccagtgaggtgttgaaggaggagcaggaaggaagtgaa 266
|||||
Db 178 GGTGGAGATGCCATCCCTGCAGTGGAGGTGTTTGAAGGGAGCCAGGAAACAAGTGAA 237
|||||
QY 267 cctggcagagctgttcaagggcaagaggtgtgctgttttgagttcctggggccttcac 326
|||||
Db 238 CCTGGCAGAGCTGTTCAAGGGCAAGAGGTGTGCTGTTTGGAGTTCCCGGGCCTTCAC 297
|||||
QY 327 cctggtatgttcaagacacacccctgcagggtttgtgagcaggtctgaggtctgaagc 386
|||||
Db 298 GCCTGGATGTTTCAAGACCCACCCTACAGGGTTTGTGGAGCAGGCTGAGGCTTGAAGGC 357
|||||
QY 387 caaggagtcagggtgtgctgtcgtgagtgatgatgccttctgtgactggcagtg 446
|||||
Db 358 CAAAGAGTCCAGGTGTGGCCCTGCTGAGTGTGTTATGATGCTTGTGACTGGCCAGTG 417
|||||
QY 447 ggcccgagcccaagcggaagcaggtgtcggtcctggctgacccactggggccctt 506
|||||
Db 418 GGCCCGAGCCCAAGCGGAAGGTTCGGCTCCTGGCTGATGCCACTGGGGCCTT 477
|||||
QY 507 tgggaaggagacagacttattactagatgctggtgtgtccatttttgggaatcagc 566
|||||
Db 478 TGGGAAGGAGACAGACTTATTACTAGATGATTCGTTAGTGTCCATCTTTGGGAATCGACG 537
|||||
QY 567 tctcaagagttctccatggtgtacagatggcatagtggaagcctgaatgtggaacc 626
|||||
Db 538 TCTCAAGAGTTCTCCATGGTGTACAGATGGCATAGTGAAGCCCTGAAATGTGGAACC 597
|||||
QY 627 agatggcacaggcctcacctgcagcctggcaccacaataatactcacagctctgaggccc 686
|||||
Db 598 AGATGACACAGGCTTCACCTGCAGTCTGGCACCACCATCATCTCACAGCTCTGAGGCC 657
|||||

```


TITLE Direct Submission
JOURNAL Submitted (13-SEP-1999) Kim I.H., Biochemistry, Paichai University,
439-6 Doma-2-Dong Seo-GU, Taejon 302-735, REPUBLIC OF KOREA

FEATURES
source 1..489
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HeLa Cell"
1..489
/gene="TPX VI"
1..489
/gene="TPX VI"
/function="peroxidase"
/codon_start=1
/evidence="experimental"
/product="human thiol peroxidase homologous protein"
/protein_id="CAB62210.1"
/db_xref="GI:6523289"
/translation="MAPIKVGDAIPAVEVEFEGEPGNKNVLAELFKGKGVLFVGPVGF
TPGCKTHLPVGEAEALKAQGVQVACLVDYFVTEGMRHAKGKRVLLADPT
GAFGKETDLLDLSLVIFGNRLKRFMSVVDGIVKALNVEPDGTLTCSLAPNIIS
OL"

BASE COUNT 107 a 114 c 163 g 105 t
ORIGIN

Query Match 60.7%; Score 489; DB 93; Length 489;
Best Local Similarity 100.0%; Pred. No. 2.1e-113;
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 atggcccaatcaagtgaggatgccatccacagcagtgagtggtttgaagggagcca 252
|||||
Db 1 ATGCCCCCAATCAAGTGGGAGATGCCATCCAGCAGTGAGGTGTTGAAGGGAGCCA 60
|||||

Qy 253 ggaacaagtgaaactggcagagctgttcaagggaaggaaggtgtgtgtttgagtt 312
|||||
Db 61 GGAACAAGGTGAACCTGGCAGAGCTGTTCAGGGCAAGAAGGTTGTGTTGGAGTT 120
|||||

Qy 313 cctggggccttccccctggatgttccaagacacacctgccagggtttgtgagcaggct 372
|||||
Db 121 CCTGGGSCCTTACCCTGGATGTTCCAGACACACCTGCCAGGGTTTGTGGAGCAGCT 180
|||||

Qy 373 gaggctgaagccaagggagtcacaggggtgtgtgtgtgtgtgtgtgtgtgtgtgt 432
|||||
Db 181 GAGGCTGAAGGCCCAAGGAGTCCAGGTGTGGCTGTCTGAGTGTATGATGCTTT 240
|||||

Qy 433 gtactgagcagtgggggcagcccaagcgaagcgaaggttcggctcctggctgat 492
|||||
Db 241 GTGACTGGGAGTGGGGCCGAGCCCAAGGCGGAAGGTTCCGGCTCCTGGCTGAT 300
|||||

Qy 493 cccactgggccccttgggaaggagacagacttattactagatgattcgtgtgtccatc 552
|||||
Db 301 CCCACTGGGCCCTTGGGAAGGAGACAGACTTATTACTAGATGATTCGGTGTCCATC 360
|||||

Qy 553 ttgggaatgcagcttcaagaggtttctccatggtgtacagatggcatagtgaagggc 612
|||||
Db 361 TTTGGGAATCGAGTCTCAAGAGGTTCTCCATGGTGGTACAGATGGCATAGTGAAGGCC 420
|||||

Qy 613 ctgaatgtgaacacagatggcagccctcactcagcctggcaccacaaatcatctca 672
|||||
Db 421 CTGAATGTGAACAGATGGCAGGCTTCCCTGAGGCTTCCCTGAGGCTTCCCTGAGGCTTCA 480
|||||

Qy 673 cagctctga 681
|||||
Db 481 CAGCTCTGA 489
|||||

RESULT 11
A99105 LOCUS
DEFINITION Sequence 9 from Patent WO9909054.
ACCESSION A99105
VERSION A99105.1 GI:6782057

KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 604)
REFERENCE
AUTHORS Falmagne,P., Wattiez,R., Bernard,A., Hermans,C. and Knoops,B.
TITLE Peroxisome-associated polypeptide, nucleotide sequence encoding
said polypeptide and their uses in the diagnosis and/or the
treatment of lung injuries and diseases, and of oxidative
stress-related disorders
JOURNAL Patent: WO 9909054-A 9 25-FEB-1999;
UNIV MONS HAINAUT (BE) ; FALMAGNE PAUL (BE) ; WATTIEZ RUDDY (BE) ;
BERNARD ALFRED (BE) ; HERMANS CEDRIC (BE) ; KNOOPS BERNARD (BE) ;
UNIV LOUVAIN (BE)
FEATURES
source 1..604
/organism="Homo sapiens"
/db_xref="taxon:9606"
161..517
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB70594.1"
/db_xref="GI:6782058"
/translation="MAPIKVGDAIPAVEVEFEGEPGNKNVLAELFKGKGVLFVGPVGF
TPGCKSVRLADPTGAFGKETDLLDLSLVIFGNRLKRFMSVVDGIVKALNVEPD
GTGLTCSLAPNIISOL"

BASE COUNT 117 a 160 c 200 g 127 t
ORIGIN

Query Match 57.4%; Score 462; DB 9; Length 604;
Best Local Similarity 82.1%; Pred. No. 1.4e-106;
Matches 604; Conservative 0; Mismatches 0; Indels 132; Gaps 1;

Qy 33 gggtagtgactagctggcgtgtgtgcccctgagacgtcagcggcgtatatactcgtgg 92
|||||
Db 1 GGGTATGGGACTAGTGGCGCTGTGGCGCTCAGAGCGCTCAGCGGGCTATATACTGTCGG 60
|||||

Qy 93 tggggccgcggtcagctcgtcggcagcgcagcagcagcgtgcagtgagagagtgaggc 152
|||||
Db 61 TGGGCGCGGCTCAGTCTGGCGCAGCGCAGCAGAGCGTGCGTGAAGGAGAGTGAGGCG 120
|||||

Qy 153 gtctggcgggggtccgcagctttcagcagcgcgcgtcagcccatgcccacatcaaggtgg 212
|||||
Db 121 GTCGTGGCGGGTCCCGAGCTTTCAGCAGAGCGCTCAGCCCATGGCCCCAATCAAGGTGG 180
|||||

Qy 213 agatgccatcccagcagctggaggtgtttgaaggggagccaggggaacagggtgaacctggc 272
|||||
Db 181 AGATGCCATCCAGCAGCTGGAGGTGTTTGAAGGGGAGCGCAGGGAACAAGGTGAACCTGGC 240
|||||

Qy 273 agagctgttcaaggcgaaggggtgtgtgtttgagttcctggggttcacccctgg 332
|||||
Db 241 AGAGCTGTTCAAGGGCAAGAAGGGTGTGTGTTGGAGTCTCTGGGGCTTTCACCCCTGG 300
|||||

Qy 333 atgttccaagacacacacctgccagggtttgtgagcaggtgaggtctgaaggccaaagg 392
|||||
Db 301 ATGTTT-----
|||||

Qy 393 agtccaggtgtgtggtctgtctgtatgttaagtgtcctttgtgactggcaggtggggcg 452
|||||
Db 307 -----
|||||

Qy 453 agcccaagcgggaagcgaaggttcgctcctgctgctatccactggggtttggaa 512
|||||
Db 307 -----CAAGGTTCCGCTCCTGGCTGATCCCTGAGGCTTGGGAA 348
|||||

Qy 513 ggagacagacttattactagatgattcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 572
|||||
Db 349 GGAGACAGACTTATTACTAGATGATTCGCTGTGTCCATCTTTGGGAATCAGCGTCTCAA 408
|||||

Qy 573 gaggttctccatgt 632
|||||

```

Db 409 GAGGTTCTCATGGTGTACAGGATGCGATAGTGAAGGCCCTGAATGTGAACACAGATGG 468
QY 633 cacaggctcacctgcagctggccacccaataatcatctacagctctgagggccctgggcc 692
Db 469 CACAGGCTCACCTGCAGCTGGCACCCATATCATCTCAGACTCAGAGCCCTGGGCC 528
QY 693 agattactctccacccctccctctatctcaccctgcccagccctgtgctggggccctgcaa 752
Db 529 AGATTACTTCTCCACCCCTCCCTATCTCACCCTGCCAGCCCTGTGCTGGGGCCCTGCAA 588
QY 753 ttggaatgttggccag 768
Db 589 TTGAATGTTGGCCAG 604

RESULT 12
AF110735 798 bp mRNA MAM 27-SEP-2000
LOCUS Sus scrofa peroxiredoxin 5 (PRDX5) mRNA, complete cds.
DEFINITION AF110735
ACCESSION AF110735
VERSION AF110735.2 GI:10305335
KEYWORDS
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 798)
AUTHORS Knoop, B., Clippe, A., Bogard, C., Arsalane, K., Wattiez, R.,
Hermans, C., Duconseille, E., Falmagne, P., and Bernard, A.
Cloning and characterization of AOE166, a novel mammalian
antioxidant enzyme of the peroxiredoxin family
J. Biol. Chem. 274 (43), 30451-30458 (1999)
99452929
PUBMED 10521424
REFERENCE 2 (bases 1 to 798)
AUTHORS Knoop, B., de Wergifosse, P. and Boone, C.
Cloning and characterization of pig AOE166/PRDX5
Unpublished
JOURNAL
REFERENCE 3 (bases 1 to 798)
AUTHORS Knoop, B., Wattiez, R., Falmagne, P., Hermans, C. and Bernard, A.
Direct Submission
JOURNAL Submitted (04-DEC-1998) Department of Biology, Catholic University
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
4 (bases 1 to 798)
AUTHORS Knoop, B., Wattiez, R., Falmagne, P., Hermans, C. and Bernard, A.
Direct Submission
JOURNAL Submitted (26-SEP-2000) Department of Biology, Catholic University
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
REMARK Sequence update by submitter
COMMENT On Sep 26, 2000 this sequence version replaced gi:10129961.
FEATURES
source
1..798
/organism="Sus scrofa"
/db_xref="taxon:9823"
/tissue_type="liver"
1..798
/gene="PRDX5"
172..660
/gene="PRDX5"
/note="AOE166; PRDX5; alkyl hydroperoxide reductase;
thioredoxin peroxidase; peroxisomal and cytosolic"
/codon_start=1
/product="peroxiredoxin 5"
/protein_id="AAG13452.2"
/db_xref="GI:10305336"
/translation="MAPIKVGDAIPSVVFEPEPKKYNLAELFKGKGVLCVPGAF
TFGCKTHLPFGVEAKLAKGIQVACLSDVDFVTWEGRAHNTGKVRLLADPT
GAFGKTELLDLDLSVLSFGNRLKRFMSVIEDGIVKSLNVEPDDTGLTCSLAPLIIS
QL"
BASE COUNT 163 a 199 c 260 g 176 t
ORIGIN

```

```

Query Match 57.1%; Score 459.8; DB 7; Length 798;
Best Local Similarity 84.2%; Pred. No. 4.8e-106;
Matches 555; Conservative 0; Mismatches 97; Indels 7; Gaps 3;

QY 145 gagtggggtctgggggggtccgcagtttcagcagagagccgctgcagcattggccccaatc 204
Db 125 GAGTGGACGTTGGCAGGGTCCGAGATTTGAGAAAGCG-CGGTGTAGCATGTGCCCCGATC 183
QY 205 aagtgtagagatgcattccagcagtgagtggtttgaagggagcagaggaacaaggtg 264
Db 184 AAGTGGGAGATGCCATCCATCCGTCGTTGGTGTGTTTGAAGGGGAGCCTGAGAAAAAGGTG 243
QY 265 aacctggcagagctgttcaaggggcagaagggtgtgtgtttggagttcctggggccctc 324
Db 244 AACCTGGCAGAGCTGTTCAAGGGCAAGAGGGGTGCTGTTTGGAGTCCCTGGGGCCCTTT 303
QY 325 acccttgatgttccaaagacacacctgcagggttttggagcaggtgaggctctgaag 384
Db 304 ACTCTGGCTGTTCCAAAGACCCACCTTCCAGGGTTCGAGCAGAGCCGAGCACTGAAG 363
QY 385 gccaaaggagtcaggtggctgtctgagtgctgttaattgattgcttcttgactggcgag 444
Db 364 GCCAAGGGGATCCAAAGTGGTGGCATGCTGAGTGTTAATGACGTCTTGTCACTGAATG 423
QY 445 tggggccgagcccaagggcgaaggcgaaggttcggctcctggctgacccactggggcc 504
Db 424 TGGGGACGACCCACACACACAGGCAAGGTCGCGCTCCTGGCTGACCCACCTGGGGCC 483
QY 505 ttgggaagagacagactattactagatgattcgtgctggtgtccatcttgggaatcga 564
Db 484 TTTGGGAAGGAGACAGATTTATTACTAGATGATTCCTTGGTGTCTCTCTTTGGGAATCGA 543
QY 565 cgtctcaaggaggttctccatggtggtacaggatggcagtagtgaggccctgaaatgtgaa 624
Db 544 CGGCTCAAGAGGTCTCCATGGTGATCGAGATGGCATAGTGAAGTCCCTGAATGTGGAG 603
QY 625 ccagatggcacaggcctcacctgcagcctggcaccacaatcatctcacagctctgagcc 684
Db 604 CCAGATGACACAGGCTCACCTGCAGCTGGCCGCCCAACATCATCTCACAGCTCTGAGGC 663
QY 685 cctg-ggceagattactctccacccctccctatctcac- - - - -tgccagccctgtg 738
Db 664 CTTGAGCCCAAGATATATTTCTTCTCCCTTTTTCAGGTGCCCGAGCTGTGGGGAGGGGGCC 723
QY 739 ctggggccctgcagtggaatggtggccagatttctgcaataaacacttgggtttgctg 797
Db 724 CGGCTCACCTCAGTCGGAACCTTGGCCAGATTCTGCAATAAACTCTTCTGTTGTCGTG 782

RESULT 13
A99104 601 bp DNA PAT 20-SEP-2000
LOCUS Sequence 8 from Patent WO9909054.
DEFINITION A99104
ACCESSION A99104
VERSION A99104.1 GI:6782055
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 601)
AUTHORS Falmagne, P., Wattiez, R., Bernard, A., Hermans, C. and Knoop, B.
Peroxisome-associated polypeptide, nucleotide sequence encoding
said polypeptide and their uses in the diagnosis and/or the
treatment of lung injuries and diseases, and of oxidative
stress-related disorders
Patent: WO 9909054-A 8 25-FEB-1999;
UNIV MONS HAINAUT (BE) ; FALMAGNE PAUL (BE) ; WATTIEZ RUDDY (BE) ;
BERNARD ALFRED (BE) ; HERMANS CEDRIC (BE) ; KNOOPS BERNARD (BE) ;
UNIV LOUVAIN (BE)
FEATURES
Location/Qualifiers
source
1..601
/organism="Homo sapiens"

```

```

CDS
/db_xref="taxon:9606"
161..514
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAB70593.1"
/db_xref="GI:6782056"
/translation="MAPIKTHLPFGVEQAEALAKAGQVQVACLSDVNDFAVFTGEMGRAH
KAGKVRLLADPTGAFGKETDILLDDSLVIFGNRLRFRSMVVDGIVKALNVEPDG
TGLTCSLAPNLSOL"

BASE COUNT      115 a      162 c      200 g      124 t
ORIGIN
33  gggtatggactagctggctgtgctgcccctgagacgtcagcgggctatatactatcg 92
Query Match      56.6%; Score 456; DB 9; Length 601;
Best Local Similarity 81.7%; Pred. No. 4.6e-105;
Matches 601; Conservative 0; Mismatches 0; Indels 135; Gaps 1;
Db 1  GGGTATGGGACTAGCTGGCTGTGCGCCCTGAGACGCTCAGCGGCTATATATCTGCGG 60
QY 93  tggggcgcggcgtcagctcgcgcagcgcgcagcaagcgtgcagtgaaagagagtgcc 152
Db 61  TGGGGCCGCGGTCAGTCTCGCGCAGCGGCGGCAAGACGGTGCAGTGAAGAGAGTGGGC 120
QY 153  gtctggcggggtccgcagtttcagcagcagcgcgtgcagccatgcccccaatcaagtg 212
Db 121  GTCTGGCGGGTCCGCGAGTTTCAGCAGAGCGGCTGCAGCCATGGCCCCAAT----- 171
QY 213  agatgccatccccagcagtgaggtgtttgaaggaggagccaggaacaagtgaaacctg 272
Db 172  -----
QY 273  agagctgttcaaggccaaggggtgtgtgtttggagttcctggggccttcacctg 332
Db 172  -----
QY 333  atgttccaaagacacacctgccagggtttgtggagcaggtgaggtcttgaaagccaa 392
Db 172  -----CAAGACACACCTGCCAGGTTTGTGGAGCAGGCTGAGGCTCTGAAGCCCAAGGG 225
QY 393  agtccaggtggcgtgctgtgagtgtaagtgtatgatgcctttgtgactggcagtgagg 452
Db 226  AGTCCAGGTGGTGGCCCTGTCTGAGTGTATGATGCTTTGTGACTGGCGAGTGGGCGC 285
QY 453  agcccaagcgaagcgaaggttcgctcctggctgattccactgggcccctttggaa 512
Db 286  AGCCCAAGCGGAAGCGAAGGTTTCGGCTCCTGGCTGATCCCACTGGGGCCCTTGGGAA 345
QY 513  ggagacagacttattactagattcgtggtgtgtccatctttgggaatcgacgtctcaa 572
Db 346  GGAGACAGACTTATTACTAGATGATTCGTTGGTGTCCATCTTTGGGAATCGACGTCTCAA 405
QY 573  gaggttccatggtgttacaggatggcatagtgaaagccctgaatgtgaaacagatg 632
Db 406  GAGGTTCTCCATGTGTGTACAGATGTCATAGTGAAGGCCCTGAAATGTGAACAGATGG 465
QY 633  cacaggcctacctgcagcctggcaccacaatacatcatcagactctgagccctggcc 692
Db 466  CACAGGCTCACCCTGCAGCCTGGGACCACAATATCATCTCAGACTCTGAGGCCCTGGCC 525
QY 693  agattacttcaccacctccctctatctcactgcccagccctgtgtgctgggcccctg 752
Db 526  AGATTACTTCCTCCACCCCTCCCTATCTACCTTCCCGAGCCCTGTGTGCTGGGCCCTG 585
QY 753  ttggaattgtggccag 768
Db 586  TTGGAATGTTGGCCAG 601

RESULT 14
AF110732
LOCUS      AF110732      837 bp      mRNA      ROD      24-OCT-1999
```


QY	249	gcaagggaacaagggtgaacctgacagctgttcacaaagggaagggtgctggtttg	308
Db	249	ACCTGGAAGAAGGTGAACCTTGGCAGAGCTGTTCAAGGACAAGAAGGTGTTTGTG	308
QY	309	agttcctggggcctcaccctggatattccaagacacacctgccagggtttgtggagca	368
Db	309	AGTCCCTGGGGCATTTACACTGGCTGTTCCAGACCATCTGCCTGGGTTGTGGAGCA	368
QY	369	ggctgaggtcctgaagcccaaggagtcagggtgtgtggcctgtctgagtgtaaatgatgc	428
Db	369	AGCGGAGCTCTGAAGGCCAAGGAGCACAAGTGGTGGCCTGTCTGAGTGTATGATGC	428
QY	429	ctttgtgactggcgagttggggccagcccaagggcggaaggccaaggttcgcctctggc	488
Db	429	CTTCGTGACTGCAGAGTGGGTCGAGCCACAGGCAAGGCAAGGTTGAGCTCTGGC	488
QY	489	tgatcccaactggggcctttgggaaggagacagacttattactagatgattcgtcgtgtc	548
Db	489	TGACCCACACTGGAGCTTTTGGAAAGGAGACAGATTTACTACTAGATGATTTTGGGTG	548
QY	549	catctttgggaatcgacgtctcaagaggttctccatctggtgtacagatggcatagtaa	608
Db	549	TCCTTTTGGGAATCGTCGGCTAAAAAGTTCTCCATGGTGATAGACAGGGCGTAGTAA	608
QY	609	ggccttgaaatggaaaccagatggcacaggccctccactgcagctggcaccccaatacat	668
Db	609	GGCACTGAACGTGGAGCGGATGGCAGAGGCTCACCTGAGCCTGGCCGCCCAACATCT	668
QY	669	ctcacagctctgagccctggcgagattactctccaccctccctatctcacc	724
Db	669	CTCACAACTCTGAGGCCCTGACCAGAAATGCTCTGACTCTCCCATCTCTCCACC	724

RESULT 15

AF305564

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AF305564

Bos taurus peroxiredoxin 5 mRNA, complete cds.

AF305564

AF305564.1

GI:12407850

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

1 (bases 1 to 881)

Leysen, G., Donnay, I. and Knoops, B.

Cloning of 4 new bovine peroxiredoxins, and screening of the complete peroxiredoxin family in different bovine tissues

Unpublished

2 (bases 1 to 881)

Leysen, G., Donnay, I. and Knoops, B.

Direct Submission

Submitted (14-SEP-2000) Biology, Universite Catholique de Louvain, Place Croix du Sud 5, Louvain-la-Neuve 1348, Belgium

Location/Qualifiers

1. .881

/organism="Bos taurus"

/db_xref="taxon:9913"

/tissue_type="liver"

77. .736

/note="thioredoxin peroxidase; alkyl hydroperoxide

CDS

BASE COUNT	177	a	231	c	276	g	197	t
ORIGIN								
Query Match	54.9%;	Score	441.6;	DB	7;	Length	881;	
Best Local Similarity	75.9%;	Pred. No.	2e-101;					
Matches	613;	Conservative	0;	Mismatches	159;	Indels	36;	Gaps
4;								
QY	34	gg	ta	g	g	a	c	t
Db	74	g	c	a	t	c	c	g
QY	94	g	g	g	g	c	c	g
Db	134	c	c	a	c	c	a	c
QY	145	g	g	g	g	c	c	g
Db	194	a	t	c	c	t	c	c
QY	199	c	c	a	a	t	c	c
Db	254	c	c	a	a	t	c	c
QY	259	a	a	g	t	a	a	c
Db	314	a	a	g	t	a	a	c
QY	319	g	c	c	t	c	c	c
Db	374	g	c	c	t	t	a	c
QY	379	c	t	a	a	g	c	c
Db	434	c	t	a	a	g	c	c
QY	439	g	c	c	a	a	g	c
Db	494	a	a	a	g	t	a	a
QY	499	g	g	g	c	c	t	t
Db	554	g	g	a	c	t	t	t
QY	559	a	a	t	c	c	a	a
Db	614	a	a	t	c	c	a	a
QY	619	g	t	a	a	c	c	a
Db	674	g	t	a	a	c	c	a
QY	679	t	a	a	c	c	c	a
Db	734	t	a	a	c	c	c	a
QY	737	t	a	a	c	c	c	a
Db	794	g	c	a	a	c	c	a
QY	778	a	t	a	a	c	c	a
Db	854	a	t	a	a	c	c	a

Search completed: October 22, 2001, 12:12:10
Job time: 5370 sec

OM of: US-09-486-167A-2 to: Issued_Patents_NA:* out_format : pfs
Date: Oct 22, 2001 10:42 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09486167/runat_22102001_064207_25045/app_query.fasta_1.222
-DB=Issued_Patents_NA -OFMT=fastap -SUFFIX=rni -GAPOP=12_000
-GAPEXT=4_000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4_500 -QGAPEXT=0.500 -XGAPOP=10.000 -XGAPEXT=0.500
-FCGAPOP=6_000 -FCGAPEXT=7_000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6_000 -DELEXT=7_000 -START=1 -MATRIX=blosum62
-TRANS=human40.cgi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM-ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09486167@cgn1_1_58 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-486-167A-2
Query length: 162
Database: Issued_Patents_NA:*
Database sequences: 324599
Database length: 94655562
Search time (sec): 57.360000

score_list:

Sequence	Strid	Orig	zScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-959-004-2 +	830.00	1877.26	6.0e-97	993	1	993
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-467-265-1 +	96.00	194.46	0.0032	918	1	918
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-467-265-1 +	96.00	194.46	0.0032	918	1	918
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-299-162A-1 +	93.50	193.48	0.0037	600	1	600
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-431-080-21 +	88.00	177.55	0.0284	807	1	807
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-431-080-31 +	88.00	177.55	0.0284	807	1	807
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-534-21 +	88.00	177.55	0.0284	807	1	807
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-938-534-31 +	88.00	177.55	0.0284	807	1	807
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-411-578-41 +	83.50	169.28	0.0820	672	1	672
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-036-987A-1 +	77.50	107.74	219.48	80161	1	80161
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-602-262-4 +	77.50	156.78	0.4069	600	1	600
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-602-262-5 +	77.50	156.78	0.4069	600	1	600
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-09-004-716-4 +	77.50	156.78	0.4069	600	1	600
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-004-716-5 +	77.50	156.78	0.4069	600	1	600
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-602-262-1 +	77.50	154.48	0.5467	737	1	737
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-602-262-3 +	77.50	154.48	0.5467	737	1	737
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-004-716-1 +	77.50	154.48	0.5467	737	1	737
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-004-716-3 +	77.50	154.48	0.5467	737	1	737
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-280-443-1 +	77.50	129.83	12.91	6671	1	6671
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-457-459-1 +	77.50	129.83	12.91	6671	1	6671
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-555-678-1 +	77.50	129.83	12.91	6671	1	6671
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-555-678-1 +	77.50	129.83	12.91	6671	1	6671
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-320-878-23 +	73.50	136.00	5.85	1693	1	1693
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-939-218A-1 +	73.50	128.93	14.48	3183	1	3183
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-940-661A-1 +	73.50	128.90	14.54	3192	1	3192
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-083-485-1 +	73.50	128.90	14.54	3192	1	3192
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-151-189-1 +	72.50	127.56	17.26	2931	1	2931
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-125-468-1 +	72.50	101.53	486.42	30001	1	30001
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-474-933-1 +	72.50	101.53	486.42	30001	1	30001
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-090-567-1 +	72.00	137.58	4.78	1081	1	1081
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-415-751-31 +	71.00	135.58	6.18	1053	1	1053
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-415-751-32 +	71.00	135.58	6.18	1053	1	1053
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-258-349-1 +	71.00	122.58	32.70	3363	1	3363
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-320-878-19 +	71.00	95.30	1.4e+03	38506	1	38506
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-594-031-65 +	70.50	147.35	1.36	332	1	332
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-5171680-4 +	70.00	141.03	3.07	527	1	527
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-883-036A-1 +	69.50	129.58	13.32	1323	1	1323
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-329-633A-1 +	69.50	126.14	20.71	1799	1	1799
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-987-123-1 +	69.00	124.04	27.11	1959	1	1959
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-987-123-3 +	69.00	124.04	27.11	1959	1	1959

/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-07-961-522-1 + 69.00 120.47 42.88 2696
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-217-438-1 + 69.00 120.47 42.88 2696
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-321-978-1 + 69.00 120.47 42.88 2696
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-710-584-1 + 69.00 120.47 42.88 2696
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-961-083-43 + 68.50 139.79 3.60 433

seq_name: /cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-959-004-2

seq_documentation_block:

Sequence 2, Application US/08959004
Patent No. 6197543
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
APPLICANT: Kasez, Matthew
TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/959,004

FILING DATE: Herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0414 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 993 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: BRAIT01

CLONE: 743725

US-08-959-004-2

alignment_scores:

Quality: 830.00 Length: 162
Ratio: 5.123 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-486-167A-2 x US-08-959-004-2 ..

Align seg 1/1 to: US-08-959-004-2 from: 1 to: 993

1 MetAlaprotilelyValGlyAspAlarleProAlaValGluValPheG1 17

|||||

385 ATGCCCCCAATCAAGTGGGAGATGCCATCCACGACGTGGGTGTTGA 434

```
17 uGlyLpProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
435 AGGGAGCCAGGCAAGTCAAGCTGCGACAGCTGTTCAAGGCAAGA 484
|||||
34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
485 AGGGTGTGCTGTTTGGAGTTCCTGGGGCTTCACCCCTGGATGTTCCAA 534
|||||
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67
|||||
535 ACACACCTGCCAGGTTGTGGAGCAGGCTGAGGCTCTGAAGCCCAAGG 584
|||||
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
585 AGTCCAGGTGGTGGCTGCTCAGTGTAAATGATGCCITTTGGACTGGCG 634
|||||
84 lUtrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
635 AGTGGGGCGGAGCCACACAGGCGAAGCAAGGTTCCGGCTCTGGCTGAT 684
|||||
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117
|||||
685 CCCATGGGGCTTTGGNAGAGACAGACTTATTACTAGATGATCGCT 734
|||||
117 uValSerLePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
|||||
735 GGTGTCCATCTTTGGAAATCGACGTCTCAAGAGGTTCTCCATGGTGTAC 784
|||||
134 InAspGlyLeValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
785 AGATGGCATAGTGAAGGCCCTTGAATGTGAACAGATGGCAGGCGCTC 834
|||||
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
|||||
835 ACCTGCAGCGCTGGCACCAATATCATCTCACAGCTC 870
|||||
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-467-265-1
seq_documentation_block:
; Sequence 1, Application US/08467265
; Patent No. 5985612
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Gentz, Reiner
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: NATURAL KILLER CELL ENHANCED FACTOR-C
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CRECHI,
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,265
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feiraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
```

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 918 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..843
; US-08-467-265-1

alignment_scores:
Quality: 96.00 Length: 153
Ratio: 1.067 Gaps: 8
Percent Similarity: 58.824 Percent Identity: 28.105

alignment_block:
US-09-486-167A-2 x US-08-467-265-1 ..
Align seg 1/1 to: US-08-467-265-1 from: 1 to: 918

2 AlaProIleLysValGlyAspAlaIleProAlaValGluValPheGluG 18
|||||
283 GCGCCCTACTTGGGAAGAACAGCT.....GTGATCGATGG 317
|||||
18 yGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysLysG 35
|||||
318 AGAA...TTTAAGGAGCTGAAGTTAAGTAT...TATCGTGGGAATACT 361
|||||
35 lYValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLysThr 51
|||||
362 TGGTTTCTCTTCTTACCCACTTGATTTTCACATTTGTGT...CCAAC 408
|||||
52 HisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysGlyVa 58
|||||
409 GAATTTATCGCTTTTGGCGACAGACTTGAAGAAATTCAGATCTATAATAC 458
|||||
68 lGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyGlu 85
|||||
459 TGAAGTGGTAGCTGC...TCGTGTGATTCACAGTTTACCCATTGGCCT 505
|||||
85 rpGlyArgAlaHisLysAlaGluGly.....LysValArg 96
|||
506 GGATTAATACCCCTCGAAGACAAGGAGGACTTGGGCCAATAAGGATCCA 555
|||||
97 LeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeuLeu 113
|||||
556 CTCTCTTTCAGATTGACCCATCAGATCTCAAGAGGACTATGGTGTATAC 605
|||||
113 uAspAspSerLeuValSerIlePheGlyAsnArgArgLeuLysArgPheS 130
|||||
606 AGAGGACTCA.....GCCACACTCTTAGAGGCTCT 637
|||||
130 erMetValValGlnAspGlyIleValLys.....AlaLeuAsnValGlu 144
|||||
638 TCATTATTGATGACAAAGGAATCTCTAAGACAAATTAATCTGAATGATCT 687
|||||
145 ProAspGly 147
|||
688 CCTGTGGGT 696

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-08-467-265-1
seq_documentation_block:
; Sequence 1, Application US/08467265
; Patent No. 6255079
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Gentz, Reiner
; APPLICANT: Rosen, Craig A.
```



```

; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-431-080-21

alignment_scores:
    Quality: 88.00      Length: 133
    Ratio: 1.086      Gaps: 6
    Percent Similarity: 60.902      Percent Identity: 25.564

alignment_block:
    US-09-486-167A-2 x US-08-431-080-21/rev ..
    Align seg 1/1 to reverse of: US-08-431-080-21 from: 1 to: 807

4 llelysalvGlyAspAlaIlePrOAlaValGluValPheGluGlyGluPr 20
   ::::||||| ||||| ::::||||| :|||
567 TTAGATAGCGGATCCTATTCCCTGATTGAGTCTTTTA...AATGAAGA 521
   :|
20 oGlyAsnlysalAsnLeuAlaGluLeuPheLysGlyLysGlyValL 37
   ::::||||| ||||| ::::||||| :|||
520 TAATGACTCTATCTCTTTGAAGAATAATCCGCGAAATAACGAGTGTGG 471
   :|
37 euPheGlyVal...ProGlyAlaPheThrProGlyCysSerLysThrHis 52
   :||| ||| ||| ||| ||||| ||||| :|||
470 TGTTTTGTGTATCCGAGGCGACGACCTGTTGTTACTAGA...CAG 424
   :|
53 LeuProGlyPheValGluGlnAlaGluLeuLysAlaLysGlyValG 69
   :||| ||||| :||| :||| :||| :||| :|||
423 GCCTGTGGATTTCTGTGACATTAACGAGGAAGTCAAG.....AA 386
   :|
69 nValValAlaCysLeuSerValAsnAspAlaPheValThrGlyGluTrpG 86
   :||| ||| :||| :||| :||| :||| :||| :||| :||| :|||
385 ATATGCTGCTCTCTTGGAGCTGAGTGCAGATCTGTGACATCCCGAAGA 336
   :|
86 LyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAspProThr 102
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
335 AGTTTCAGAGTAACAAATTTGCCATATCATTTACTAAGCGATCCCAAG 286
   :|
103 GlyAlaPhe.....GlyLysGluThrAspLeu.....Le 112
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
285 AGAGAGTTTATTGGTGTCTAGGAGCGCAAAAAACGCCACTTCTGGTTC 236
   :|
112 uLeuAspAspSerLeuValSerIlePheGlyAsnArgLeuLysArg 128
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
235 TATTATGCGCATTTTCATTTTGTGTGGGAGGTTAAATTCATAAAGA 187

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-431-080-31

seq_documentation_block:
; Sequence 31, Application US/08431080
; Patent No. 5698686
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; APPLICANT: Singer, Miriam S.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433

```

```

; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: ARCD:155/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-431-080-21

alignment_scores:
    Quality: 88.00      Length: 133
    Ratio: 1.086      Gaps: 6
    Percent Similarity: 60.902      Percent Identity: 25.564

alignment_block:
    US-09-486-167A-2 x US-08-431-080-21/rev ..

    Align seg 1/1 to reverse of: US-08-431-080-21 from: 1 to: 807

        4 llelysalvGlyAspAlaIlePrOAlaValGluValPheGluGlyGluPr 20
          ::::||||| ||||| ::::||||| ::::||||| ::::|||||
        567 TTAGATAGCGGATCCCTATTCCCTGATTGAGTCTTTTA...AATGAAGA 521

        20 oGlyAsnlysalAsnLeuAlaGluLeuPheLysGlyLysGlyValL 37
          ::::||||| ||||| ::::||||| ::::||||| ::::|||||
        520 TAATGACTCTATCTCTTTGAAGAATAATCCGCGAAATAACAGATGTGG 471

        37 euPheGlyVal...ProGlyAlaPheThrProGlyCysSerLysThrHis 52
          :::| ||| ||| ||| ||| ||||| ||||| :::: ||| ::::
        470 TGTTTTGTGTATCCAGGCGACGACGCTGTTGTTACTAGA...CAG 424

        53 LeuProGlyPheValGluGlnAlaGluLeuLysAlaLysGlyValG 69
          ||||| :::: ||| :::: ||||| :::: ||||| ::::
        423 GCCTGTGGATTTCTGTGACAATTACCGAGAACTCAAG.....AA 386

        69 nValValAlaCysLeuSerValAsnAspAlaPheValThrGlyGluTrpG 86
          :::||| ::::||||| ::::||||| ::::||||| ::::|||||
        385 ATATGCTGCTCTCTTGGAGCTGAGTGCAGATCTGTGACATCCCAAGAA 336

        86 LyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAspProThr 102
          ||| :::: ||| :::: ||||| :::: ||||| :::: |||||
        335 AGTTTCAGAGTAACAAATTTGCCATATCATTTTACTAAGCGATCCCAAG 286

        103 GlyAlaPhe.....GlyLysGluThrAspLeu.....Le 112
          ||| :::: ||| :::: ||||| :::: ||||| :::: |||||
        285 AGAGAGTTTATTGGTGTCTAGGAGCCAAAAACGCCACTTCTCTGGTTC 236

        112 uLeuAspAspSerLeuValSerIlePheGlyAsnArgLeuLysArg 128
          ::: :::: ||||| :::: ||||| :::: ||||| :::: |||||
        235 TATTATGCGCATTTTCATTTTGTGGTGGGAAGTTAAATTCAAAAGA 187

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-431-080-31

seq_documentation_block:
; Sequence 31, Application US/08431080
; Patent No. 5698686
; GENERAL INFORMATION:
; APPLICANT: Gottschling, Daniel E.
; APPLICANT: Singer, Miriam S.
; TITLE OF INVENTION: Telomerase Compositions and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433

```



```
3  ProIleIysValGlyAspAlaIleProAlaValGluValPheGluGlyG1 19
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
4  CCCTTGAACCTGGAGATTCTTTCAGACTTCCAGCG...GAGCGCT 50
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
19 uProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysGlyV 36
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
51 GGCGCGGAGCACTTCGGTTGCACAGTACTTGGGGGACAGTGGGAG 100
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
36 aLeuPheGlyValProGlyAlaPheThrProGlyCysSerLysThrHis 52
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
101 TGATGTTACGCCACCGAAGCACTTCCATCCCGTTGCACA...ACGGAG 147
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
53 LeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaGlyValG1 69
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
148 CTCGCCGAGCCGCTGAAGCTCCAGGCGATCTTCCAGGAAGAACTGCA 197
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
69 nValValAlaCysLeuSerValAsnAspAlaPheValThrGlyGluTIP 86
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
198 ACTCGTTGGC...TTCCTCTGCACGACCTGCAGAGCCACAGATGGG 244
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
86 lyArgAlaHisLysAla.....GluGlyLysVal.....Arg 96
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
245 CGAAGATATAATGGCTATGCAGGCGGATCTGGGAACCTTGGCGTTCC 294
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
97 LeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeuLe 113
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
295 CTCGTTGGGACCC.....AATAGGGAACCTG..... 309
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
113 uAspAspSerLeuValSerIlePheGlyAsnArgArgLeuLysArgPhe 130
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
310 .....AATAGGGAACCTG..... 321
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
130 erMetValValGlnAspGlyIleValLysAlaLeuAsnValGluPro 146
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
322 .....GCCGCGAGTTTGGGAATATGATCCTGCAGAAAAGGACAAA 366
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
147 GlyThrGlyLeuThrCys 152
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
367 GGGCTGCTTGTGACTTGC 384
```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-036-987A-1

```
seq_documentation_block:
; Sequence 1, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow AgroSciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-NAR-1998
; CLASSIFICATION: 435
```

```
ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80161 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-036-987A-1

alignment_scores:
Quality: 80.00 Length: 175
Ratio: 0.930 Gaps: 12
Percent Similarity: 49.143 Percent Identity: 29.714

alignment_block:
US-09-486-167A-2 x US-09-036-987A-1 ..
Align seg 1/1 to: US-09-036-987A-1 from: 1 to: 80161

12 AlaValGluValPheGluGlyGluProGlyAsnLysValAsnLeuAlaG1 28
||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
65551 GCGCGGAGGGTTTCGAGGGTTATCTAGTAAATGCG...ACGCGGGGAGG 65597
||||| ||||| ||||| ||||| ||||| ||||| |||||

28 uLeuPheLysGlyLysLysGlyValLeuPheGlyValProGly..... 42
||||| ||||| ||||| ||||| ||||| ||||| |||||

65598 CGTGTTCGGGTGGGTTCGTTGTTTCGTTGTTTCGAGGCTCTCGCG 65647
||||| ||||| ||||| ||||| ||||| ||||| |||||

43 ..AlaPheThrProGlyCysSerLysThr.....HisLeuPro 54
||||| ||||| ||||| ||||| ||||| ||||| |||||

65648 TGACGGTGGATACGGCGGTTCGTCGTTGGTGGCGCTGCACCTGGCG 65697
||||| ||||| ||||| ||||| ||||| ||||| |||||

65698 GGT.....CAAGCACTCGGCTCTGTTGAGTGTGATCTGGC 65732
||||| ||||| ||||| ||||| ||||| ||||| |||||

65 ..AlaLysGlyValGlnValValAlaCysLeuSerValAsnAspAlap 80
||||| ||||| ||||| ||||| ||||| ||||| |||||

65733 TCTTCGGGTGGTGTGACGGTGTGATGGCCACGCGGGATG.....T 65773
||||| ||||| ||||| ||||| ||||| ||||| |||||

80 heValThrGlyGluTrpGlyArgAlaHisLysAlaGluGlyLysValArg 96
||||| ||||| ||||| ||||| ||||| ||||| |||||

65774 TCGTGGAGTTTTCGGCTCAACGGGCTTGGCGGATGGCGGTGCAAG 65823
||||| ||||| ||||| ||||| ||||| ||||| |||||

97 .....LeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAs 110
||||| ||||| ||||| ||||| ||||| ||||| |||||

65824 TCGTTTCGGCGGCTGCGGATGGCACCGGT...TGGGAGAGGCGCGG 65870
||||| ||||| ||||| ||||| ||||| ||||| |||||

110 pLeuLeuLeuAspSerLeuValSerIlePheGlyAsnArgLeuL 127
||||| ||||| ||||| ||||| ||||| ||||| |||||

65871 CTGTGTGTG.....TTGG 65884
||||| ||||| ||||| ||||| ||||| ||||| |||||

127 ysArgPheSerMetValValGlnAspGly.....IleVal 138
||||| ||||| ||||| ||||| ||||| ||||| |||||

65885 AGCGGCTGCGGATGCGCGGCAACGGGACACGGGTTCTGGCGGTCTG 65934
||||| ||||| ||||| ||||| ||||| ||||| |||||

139 LysAlaLeuAsnValGluProAspGly.....ThrGlyLeuThrCys 153
||||| ||||| ||||| ||||| ||||| ||||| |||||

65935 CGGGGTACGGGTGAATCAGATGGTCCGTCGAATGTTTGACG..... 65979
||||| ||||| ||||| ||||| ||||| ||||| |||||

153 rLeuAlaProAsnIleLeuSerGln 161
||||| ||||| ||||| ||||| ||||| ||||| |||||

65980 ....GCGCGGAATGGGCCCTCGCAG 66000
||||| ||||| ||||| ||||| ||||| ||||| |||||

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-602-262-4
```



```
582 AAAGCATTCATTGGTCAACCGCCCTTAATTTCAAACAACACAGCGTTGT 533
22 .....AsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
   : : : : : : : : : : : : : : : : : : : : : : : :
532 GAATGCGGATTTCAAGGAATTTCACTTTGTGAG...TTCAAGGAAAT 486
   : : : : : : : : : : : : : : : : : : : : : : : :
34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
   : : : : : : : : : : : : : : : : : : : : : : : :
485 ATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 439
   : : : : : : : : : : : : : : : : : : : : : : : :
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67
   : : : : : : : : : : : : : : : : : : : : : : : :
438 ACAGAGATAATTCCTTTTCTGCTGATGTCGAGTTCAAAAAATTAGA 389
   : : : : : : : : : : : : : : : : : : : : : : : :
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
   : : : : : : : : : : : : : : : : : : : : : : : :
388 TGTAGCTGTTATGGCATGC...TCACTGATTCACATTTTTCACACCTTG 342
   : : : : : : : : : : : : : : : : : : : : : : : :
84 luTrpGlyArgAlaHisLysAlaGluGly.....LysVal 95
   : : : : : : : : : : : : : : : : : : : : : : : :
341 CATGGGTAATACCGACCGAAAAATGGGTGAGTCCGTCAGATGAATATA 292
   : : : : : : : : : : : : : : : : : : : : : : : :
96 ArgLeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeu 112
   : : : : : : : : : : : : : : : : : : : : : : : :
291 CCAATTCCTGCTGATACCAATCATACATATTAGTAGGGCATATGGCGTG 242
   : : : : : : : : : : : : : : : : : : : : : : : :
112 uLeuAspAspSerLeuValSerIlePheGlyAsnArgArgLeuLysArg 129
   : : : : : : : : : : : : : : : : : : : : : : : :
241 CAAGGAAGATGATGGCATGCTTACCGTGA.....T 210
   : : : : : : : : : : : : : : : : : : : : : : : :
129 heSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
   : : : : : : : : : : : : : : : : : : : : : : : :
209 TATTATCATTTGATCCAAAAGGATTTTGGCAGAAATCACAATC 166
   : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-004-716-4
seq_documentation_block:
; Sequence 4, Application US/09004716
; Patent No. 6031077
; GENERAL INFORMATION:
; APPLICANT: Klimowski, Laura
; APPLICANT: Tripp, Cynthia A.
; TITLE OF INVENTION: PARASITIC HELMINTH LARVAL THIOL SPECIFIC
; TITLE OF INVENTION: ANTIOXIDANT PROTEINS, NUCLEIC ACID MOLECULES, AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/004,716
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,262
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Conneil, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-45
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 4:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-004-716-4

alignment_scores:
  Quality: 77.50      Length: 148
  Ratio: 0.994      Gaps: 6
  Percent Similarity: 52.703      Percent Identity: 22.973

alignment_block:
US-09-486-167A-2 x US-09-004-716-4 ..
Align seg 1/1 to: US-09-004-716-4 from: 1 to: 600

14 GluValPheGluGlyGluProGly..... 21
   : : : : : : : : : : : : : : : : : :
19 AAAGCATTCATTGGTCAACCGCCCTTAATTTCAAACAACACAGCGTTGT 68
   : : : : : : : : : : : : : : : : : :
22 .....AsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
   : : : : : : : : : : : : : : : : : :
69 GAATGCGGATTTCAAGGAATTTCACTTTGTGAG...TTCAAGGAAAT 115
   : : : : : : : : : : : : : : : : : :
34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
   : : : : : : : : : : : : : : : : : :
116 ATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 162
   : : : : : : : : : : : : : : : : : :
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67
   : : : : : : : : : : : : : : : : : :
163 ACAGAGATAATTCCTTTTCTGCTGATGTCGAGTTCAAAAAATTAGA 212
   : : : : : : : : : : : : : : : : : :
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
   : : : : : : : : : : : : : : : : : :
213 TGTAGCTGTTATGGCATGC...TCACTGATTCACATTTTTCACACCTTG 259
   : : : : : : : : : : : : : : : : : :
84 luTrpGlyArgAlaHisLysAlaGluGly.....LysVal 95
   : : : : : : : : : : : : : : : : : :
260 CATGGTAATACCGACCGAAAAATGGGTGAGTCCGTCAGATGAATATA 309
   : : : : : : : : : : : : : : : : : :
96 ArgLeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeu 112
   : : : : : : : : : : : : : : : : : :
310 CCAATTCCTGCTGATACCAATCATACATATTAGTAGGGCATATGGCGTG 359
   : : : : : : : : : : : : : : : : : :
112 uLeuAspAspSerLeuValSerIlePheGlyAsnArgArgLeuLysArg 129
   : : : : : : : : : : : : : : : : : :
360 CAAGGAAGATGATGGCATGCTTACCGTGA.....T 391
   : : : : : : : : : : : : : : : : : :
129 heSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
   : : : : : : : : : : : : : : : : : :
392 TATTATCATTTGATCCAAAAGGATTTTGGCAGAAATCACAATC 435
   : : : : : : : : : : : : : : : : : :
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-004-716-5
seq_documentation_block:
; Sequence 5, Application US/09004716
; Patent No. 6031077
; GENERAL INFORMATION:
; APPLICANT: Klimowski, Laura
; APPLICANT: Tripp, Cynthia A.
; TITLE OF INVENTION: PARASITIC HELMINTH LARVAL THIOL SPECIFIC
; TITLE OF INVENTION: ANTIOXIDANT PROTEINS, NUCLEIC ACID MOLECULES, AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
```

```
;;
;; ZIP: 80203
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/004,716
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/602,262
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Connell, Gary J.
;; REGISTRATION NUMBER: 32,020
;; REFERENCE/DOCKET NUMBER: 2618-45
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 303/863-9700
;; TELEFAX: 303/863-0223
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 600 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cdna
US-09-004-716-5

alignment_scores:
  Quality: 77.50      Length: 148
  Ratio: 0.994       Gaps: 6
  Percent Similarity: 52.703   Percent Identity: 22.973

alignment_block:
US-09-486-167A-2 x US-09-004-716-5 rev ..

Align seg 1/1 to reverse of: US-09-004-716-5 from: 1 to: 600

14 GluValPheGluGluProGly.....21
   ::::: ||| :::::
582 AAAGCATTGTTGGTCAACCGGCCCTAATTTCAAACAACACCGGTTGT 533
22 .....AsnLysValasnLeuAlaGluLeuPheLysGlyLysL 34
   ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
532 GAATGGCGATTTCAGGAAATTTCACTTTGTCAG...TTCAAAGGAAAT 486
34 ysglyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
485 ATGTGTCCTCTCTTTTATCCTCTCGATTTCACTTTCGTTTGC...CCA 439

51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67
   ||| ::::: ::::: ||| ::::: ||| ::::: ||| ::::: |||
438 ACAGAGATAATTCGTTTCTGATCGTATGCGGAGTTCAAATAATAGA 389
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
388 TGTAGCTGTATGGCATGC...TCAACTGATTCACATTTTTCACACCTTG 342
84 LutrPcGlyArgAlaHisLysAlaGluGly.....LysVal 95
   ||| ::::: ::::: |||
341 CATGGTAATACCGACCGCAAAATGGGTGGACTCGGTGAGATGATATA 292
96 ArgLeuLeuAlaAspProThrGlyAlaPheGlyLysGluThrAspLeu 112
   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
291 CCAATTCCTGCTGATACCAATCATCAATAGTAGGCATATAGGCGTGTG 242
112 uLeuAspAspSerLeuValSerIlePheGlyAsnArgArgLeuLysArg 129
   , | ::::: |||
241 CAAGGAAGATGATGGCATGCTTACCGTGA.....T 210
129 heSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
```

```
;;
;; seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-602-262-1
;;
;; seq_documentation_block:
;; Sequence 1, Application US/08602262
;; Patent No. 5744593
;; GENERAL INFORMATION:
;; APPLICANT: Klimowski, Laura
;; APPLICANT: Tripp, Cynthia A.
;; TITLE OF INVENTION: PARASITIC HELMINTH LARVAL THIOL SPECIFIC
;; TITLE OF INVENTION: ANTIOXIDANT PROTEINS, NUCLEIC ACID MOLECULES, AND USES
;; TITLE OF INVENTION: THEREOF
;; NUMBER OF SEQUENCES: 10
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sheridan Ross & McIntosh
;; STREET: 1700 Lincoln Street, Suite 3500
;; CITY: Denver
;; STATE: CO
;; COUNTRY: U.S.A.
;; ZIP: 80203
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/602,262
;; FILING DATE:
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Connell, Gary J.
;; REGISTRATION NUMBER: 32,020
;; REFERENCE/DOCKET NUMBER: 2618-45
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 303/863-9700
;; TELEFAX: 303/863-0223
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 737 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cdna
;; NAME/KEY: CDS
;; LOCATION: 6..605
US-08-602-262-1

alignment_scores:
  Quality: 77.50      Length: 148
  Ratio: 0.994       Gaps: 6
  Percent Similarity: 52.703   Percent Identity: 22.973

alignment_block:
US-09-486-167A-2 x US-08-602-262-1 ..

Align seg 1/1 to: US-08-602-262-1 from: 1 to: 737

14 GluValPheGluGluProGly.....21
   ::::: ||| ::::: ||| ::::: ||| ::::: |||
24 AAAGCATTGTTGGTCAACCGGCCCTAATTTCAAACAACACCGGTTGT 73
22 .....AsnLysValasnLeuAlaGluLeuPheLysGlyLysL 34
   ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
74 GAATGGCGATTTCAGGAAATTTCACTTTGTCAG...TTCAAAGGAAAT 120
34 ysglyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 ATGTGTCCTCTCTTTTATCCTCTCGATTTCACTTTCGTTTGC...CCA 167
```

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2001, 12:15:26 ; Search time 58.12 Seconds
(without alignments)
2622.083 Million cell updates/sec

Title: US-09-486-167A-1
 Perfect score: 805
 Sequence: 1 accacagagcgagtggaag.....ttatggtttcgcgaaaaaa 805

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

searched: 324599 seqs, 94655562 residues

Word size : 0

Total number of hits satisfying chosen parameters: 649198

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Listing first 45 summaries

```
Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	798	99.1	993	4	US-08-959-7043-2	Sequence 2, Appli
C 2	17	2.1	37	1	US-08-383-743A-19	Sequence 19, Appl
C 3	17	2.1	37	3	US-08-808-881-19	Sequence 19, Appl
C 4	17	2.1	37	3	US-09-017-631-19	Sequence 19, Appl
C 5	17	2.1	37	5	PCT-US93-071116-19	Sequence 19, Appl
C 6	17	2.1	1280	4	US-09-060-756-4	Sequence 4, Appli
C 7	17	2.1	1930	1	US-07-982-113-1	Sequence 1, Appli
C 8	16	2.0	44	1	US-08-348-143-12	Sequence 12, Appl
C 9	16	2.0	44	1	US-08-571-785-12	Sequence 12, Appl
10	16	2.0	270	1	US-08-143-576-2	Sequence 2, Appli
C 11	16	2.0	439	4	US-09-060-756-240	Sequence 240, App
C 12	16	2.0	492	5	PCT-US95-08295-9	Sequence 9, Appli
C 13	16	2.0	558	4	US-09-446-504-69	Sequence 69, Appl
14	16	2.0	732	1	US-08-145-006C-8	Sequence 8, Appli
15	16	2.0	732	5	PCT-US94-00545-8	Sequence 8, Appli
16	16	2.0	794	3	US-09-010-809-14	Sequence 14, Appl
C 17	16	2.0	1084	2	US-08-184-009-110	Sequence 110, App
C 18	16	2.0	1084	2	US-08-458-356-110	Sequence 110, App
C 19	16	2.0	1094	2	US-08-184-009-109	Sequence 109, App
C 20	16	2.0	1094	2	US-08-458-356-109	Sequence 109, App
21	16	2.0	1142	4	US-08-768-373-5	Sequence 5, Appli
C 22	16	2.0	1353	2	US-08-712-939B-2	Sequence 2, Appli
C 23	16	2.0	1474	5	PCT-US94-00545-19	Sequence 19, Appl
C 24	16	2.0	1691	2	US-08-993-118-8	Sequence 8, Appli
C 25	16	2.0	1691	3	US-08-845-528C-8	Sequence 8, Appli
C 26	16	2.0	1800	2	US-08-733-825-1	Sequence 1, Appli
27	16	2.0	1896	4	US-09-345-468-24	Sequence 24, Appl

RESULT 1
US-08-959-004-2
Sequence 2, Application US/08959004
Patent NO. 6197543
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
APPLICANT: Kaser, Matthew
TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/959,004
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0414 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 993 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITUT01
CLONE: 743725

US-08-959-004-2

Query Match 99.1%; Score 798; DB 4; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 798; Conservative 0; Mismatches 0; Indels

[illegible]

RESULT 2

RESOLUTION 2
US-08-383-743A-19/c
; Sequence 19, Application US/08383743A
; Patent No. 5614184
; GENERAL INFORMATION:
; APPLICANT: SYKOWSKI, Arthur J.
; * APPLICANT: Grodberg, Jennifer
; TITLE OF INVENTION: RECOMBINANT HUMAN
; TITLE OF INVENTION: ALTERED BIOLOGICAL

NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/383,743A
FILING DATE: 08/383,743
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/113,080
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/920,810
FILING DATE: 28-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/920,810
FILING DATE: 28-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: NEDH92-04A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-383-743A-19

Query Match 2.1%; Score 17; DB 1; Length 37;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 17; Conservative 0; Mismatches 0; Indels

Qy 372 tgaggctctgaagcca 388
| | | | | | | |
Db 23 TGAGGCTCTGAAGGCCA 7

RESULT 3

```

RESOL. 3
US-08-808-881-19/c
; Sequence 19, Application US/08080881A
; Patent No. 6048971
; GENERAL INFORMATION:
; APPLICANT: Sytkowski, Arthur J.
; APPLICANT: Grodberg, Jennifer
; TITLE OF INVENTION: RECOMBINANT HUMAN ERYTHROPOIETIN WITH
; TITLE OF INVENTION: ALTERED BIOLOGICAL ACTIVITY
; FILE REFERENCE: NE092-04A22
; CURRENT APPLICATION NUMBER: US/08/808,881A
; CURRENT FILING DATE: 1997-02-28
; EARLIER APPLICATION NUMBER: US 08/383,743
; EARLIER FILING DATE: 1995-02-02
; EARLIER APPLICATION NUMBER: US 08/113,080
; EARLIER FILING DATE: 1993-08-26
; EARLIER APPLICATION NUMBER: US 07/920,810
; EARLIER FILING DATE: 1992-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
;

```

; LENGTH: 37
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide encoding protein mutant.
US-08-808-881-19

Query Match 2.1%; Score 17; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 372 tgaggctctgaagccca 388
|||||
Db 23 TGAGGCTCTGAAGGCCA 7

RESULT 4
US-09-017-631-19/c
; Sequence 19, Application US/09017631B
; Patent No. 6153407
; GENERAL INFORMATION:
; APPLICANT: Sytkowski, Arthur J.
; APPLICANT: Grodberg, Jennifer
; TITLE OF INVENTION: ERYTHROPOIETIN WITH ALTERED BIOLOGICAL
; TITLE OF INVENTION: ACTIVITY
; FILE REFERENCE: NEDH92-04A22A
; CURRENT APPLICATION NUMBER: US/09/017,631B
; CURRENT FILING DATE: 1998-02-03
; EARLIER APPLICATION NUMBER: US 08/808,881
; EARLIER FILING DATE: 1997-02-28
; EARLIER APPLICATION NUMBER: US 08/383,743
; EARLIER FILING DATE: 1995-02-02
; EARLIER APPLICATION NUMBER: US 08/113,080
; EARLIER FILING DATE: 1993-08-26
; EARLIER APPLICATION NUMBER: US 07/920,810
; EARLIER FILING DATE: 1992-07-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic oligonucleotide encoding protein mutant.
US-09-017-631-19

Query Match 2.1%; Score 17; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 372 tgaggctctgaagccca 388
|||||
Db 23 TGAGGCTCTGAAGGCCA 7

RESULT 5
PCT-US93-07116-19/c
; Sequence 19, Application PC/TUS9307116
; GENERAL INFORMATION:
; APPLICANT: New England Deaconess Hospital
; TITLE OF INVENTION: RECOMBINANT HUMAN ERYTHROPOIETIN WITH
; TITLE OF INVENTION: ALTERED BIOLOGICAL ACTIVITY
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07116
; FILING DATE: 19930728
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/920,810
; FILING DATE: 28-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: NEDH92-04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PCT-US93-07116-19

Query Match 2.1%; Score 17; DB 5; Length 37;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 372 tgaggctctgaagccca 388
|||||
Db 23 TGAGGCTCTGAAGGCCA 7

RESULT 6
US-09-060-756-4
; Sequence 4, Application US/09060756
; Patent No. 6183957
; GENERAL INFORMATION:
; APPLICANT: Cole, Stewart
; APPLICANT: Buchrieser-Brosch, Roland
; APPLICANT: Gordon, Stephen
; APPLICANT: Billault, Alain
; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
; TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
; TITLE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA
; FILE REFERENCE: 3495-0169
; CURRENT APPLICATION NUMBER: US/09/060,756
; CURRENT FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1280
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-09-060-756-4

Query Match 2.1%; Score 17; DB 4; Length 1280;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 gtcggtggggccggcgg 104
|||||
Db 679 gtcggtggggccggcgg 695

RESULT 7
US-07-982-112-1/c
; Sequence 1, Application US/07982112

; Patent No. 5346813
; GENERAL INFORMATION:
; APPLICANT: BODENMULLER, Heinz
; APPLICANT: DESSAUER, Andreas
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF A SMALL
; TITLE OF INVENTION: CELL LUNG CARCINOMA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikolaus, Marmelstein, Murray & Oram
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/982,112
; FILING DATE: 19921125
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/617,102
; FILING DATE: 21-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Chin, Monica F.
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: 564-1906
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1930 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10..933
; US-07-982-112-1

Query Match 2.18; Score 17; DB 1; Length 1930;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 333 atgttccaagacacc 349
|||||
Db 1668 ATGTTCCAAGACACC 1652

RESULT 8
US-08-348-143-12/c
; Sequence 12, Application US/08348143
; Patent No. 5506205
; GENERAL INFORMATION:
; APPLICANT: TAJIMA, HISAO
; APPLICANT: KITAGAWA, KOICHIRO
; APPLICANT: OHNO, HIROYUKI
; APPLICANT: UENO, TOSHIO
; TITLE OF INVENTION: A No. 5506205el Polypeptide of Protein p140 and DNAs
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.

; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/348,143
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 315806/1993
; FILING DATE: 24-NOV-1993
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (synthetic)
; US-08-348-143-12

Query Match 2.04; Score 16; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 575 ggtttccatggtggt 590
|||||
Db 26 GGTTCCTCATGTGT 11

RESULT 9
US-08-571-785-12/c
; Sequence 12, Application US/08571785
; Patent No. 5804411
; GENERAL INFORMATION:
; APPLICANT: TAJIMA, HISAO
; APPLICANT: KITAGAWA, KOICHIRO
; APPLICANT: OHNO, HIROYUKI
; APPLICANT: UENO, TOSHIO
; TITLE OF INVENTION: A No. 5804411el Polypeptide of Protein p140
; TITLE OF INVENTION: and DNAs encoding it
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/571,785
; FILING DATE: 13-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/348,143
; FILING DATE: 23-NOV-1994
; APPLICATION NUMBER: JP 315806/1993
; FILING DATE: 24-NOV-1993
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)293-7060
; TELEFAX: (202)293-7860
; TELEX: 6491103

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
US-08-571-785-12

Query Match 2.0%; Score 16; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 575 ggttctccatggtggt 590
|||||

Db 26 GGTTCTCCATGTTGTT 11

RESULT 10

US-08-143-576-2
Sequence 2; Application US/08143576
Patent No. 5643761

GENERAL INFORMATION:

APPLICANT: Fisher, Paul B.
APPLICANT: Jiang, Hongping
TITLE OF INVENTION: METHOD FOR GENERATING A SUBSTRUCTED
TITLE OF INVENTION: CDNA LIBRARY AND USES OF THE GENERATED LIBRARY
NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: John P. White, c/o Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/143,576
FILING DATE: 25-OCT-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 43563/JPW/AKC
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 270 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-143-576-2

Query Match 2.0%; Score 16; DB 1; Length 270;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 641 tcaactgcagctgpc 656
|||||

Db 121 TCACCTGCAGCTGCC 136

RESULT 11

US-09-060-756-240/c
Sequence 240; Application US/09060756
Patent No. 6183957

GENERAL INFORMATION:

APPLICANT: Cole, Stewart
APPLICANT: Buchrieser-Brosch, Roland
APPLICANT: Gordon, Stephen
APPLICANT: Billault, Alain
TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
TITLE OF INVENTION: LIBRARY APPLICATION TO THE DETECTION OF MYCOBACTERIA
FILE REFERENCE: 3495-0169
CURRENT APPLICATION NUMBER: US/09/060.756
CURRENT FILING DATE: 1998-04-16
NUMBER OF SEQ ID NOS: 743
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 240
LENGTH: 439
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
US-09-060-756-240

Query Match 2.0%; Score 16; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 681 aggcctggggccagat 696
|||||

Db 184 AGGCCCTGGGCCAGAT 169

RESULT 12

PCT-US95-08295-9
Sequence 9; Application PC/TUS9508295
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: BREAST SPECIFIC GENES AND PROTEINS
NUMBER OF SEQUENCES: 30
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08295
FILING DATE: 30-JUN-1995

CLASSIFICATION:

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 492 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2..490
PCT-US95-08295-9

Query Match 2.0%; Score 16; DB 5; Length 492;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 tggagttctctggggcc 321
|||||

Db 184 TGGAGTTCTCTGGGGCC 199

RESULT 13

US-09-446-504-69/c
; Sequence 69, Application US/09446504
; Patent No. 6218150
; GENERAL INFORMATION:
; APPLICANT: UEMORI, Takashi
; APPLICANT: SATO, Yoshimi
; APPLICANT: FUJITA, Tomoko
; APPLICANT: MIYAKE, Kazuo
; APPLICANT: MURAI, Hiroyuki
; APPLICANT: ASADA, Kiyozi
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
; FILE REFERENCE: 1422-408PCT
; CURRENT APPLICATION NUMBER: US/09/446,504
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: PCT/JP98/02845
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: JP 9-187496
; PRIOR FILING DATE: 1997-06-26
; PRIOR APPLICATION NUMBER: JP 9-320692
; PRIOR FILING DATE: 1997-11-27
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 69
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Pyrococcus furiosus
US-09-446-504-69

Query Match 2.0%; Score 16; DB 4; Length 558;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 569 tcaagaggttctccat 584
|||||

Db 332 TCAAGAGGTTCTCCAT 317

RESULT 14
US-08-145-006C-8
; Sequence 8, Application US/08145006C
; Patent No. 5656452
; GENERAL INFORMATION:
; APPLICANT: Rao, Anjana
; APPLICANT: Hogan, Patrick Gerald
; APPLICANT: McCaffrey, Patricia
; APPLICANT: Jain, Jugnu
; TITLE OF INVENTION: NF-ATp, A T LYMPHOCYTE
; TITLE OF INVENTION: DNA-BINDING PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,006C
; FILING DATE: October 29, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/017,052
; FILING DATE: February 11, 1993
; APPLICATION NUMBER: 08/006,067
; FILING DATE: January 15, 1993
; ATTORNEY/AGENT INFORMATION:

; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 04590/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-145-006C-8

Query Match 2.0%; Score 16; DB 1; Length 732;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 306 tggagttctggggcc 321
|||||

Db 532 TGGAGTTCTGGGGCC 547

RESULT 15
PCT-US94-00545-8
; Sequence 8, Application PC/TUS9400545
; GENERAL INFORMATION:
; APPLICANT: Rao, Anjana
; APPLICANT: Hogan, Patrick Gerald
; APPLICANT: McCaffrey, Patricia
; APPLICANT: Jain, Jugnu
; TITLE OF INVENTION: NF-ATp, A T LYMPHOCYTE
; TITLE OF INVENTION: DNA-BINDING PROTEIN
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00545
; FILING DATE: 18-JAN-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/145,006
; FILING DATE: October 29, 1993
; APPLICATION NUMBER: 08/017,052
; FILING DATE: February 11, 1993
; APPLICATION NUMBER: 08/006,067
; FILING DATE: January 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 04590/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

us-09-486-167a-1.l.oligo.rni

Mon Oct 22 14:46:47 2001

PCT-US94-00545-8

Query Match 2.0%; Score 16; DB 5; Length 732;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0; .

Qy 306 tggagttcctggggcc 321
 |||||
 Db 532 TGGAGTTCCTGGGGCC 547

Search completed: October 22, 2001, 13:44:27
 Job time: 5341 sec

Mon Oct 22 14:46:45 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: October 22, 2001, 12:13:50 ; Search time 1307.97 Seconds
(without alignments)
9519.743 Million cell updates/sec

Title: US-09-486-167A-1
Perfect score: 805
Sequence: 1 gccagggagggagtggaag.....ttgtgttgcggaaaaaaa 805

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues

Word size : 0
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_om.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_pl1.*
13: gb_pl2.*
14: gb_pl3.*
15: gb_pl4.*
16: gb_ba1.*
17: em_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htg_hum1.*
23: em_htg_hum2.*
24: em_htg_hum3.*
25: em_htg_hum4.*
26: em_htg_hum5.*
27: em_htg_hum6.*
28: em_htg_hum7.*
29: em_htg_hum8.*
30: em_htg_inv1.*
31: em_htg_inv2.*
32: em_htg_other.*
33: em_htg_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vl.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vl1.*
59: gb_vl2.*
60: gb_htg1.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_rol.*
95: gb_rol2.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	805	100.0	805	9 A99097	A99097 Sequence 1
2	805	100.0	805	88 AF110731	AF110731 Homo sapi
3	785	97.5	875	89 AF242525	AF242525 Homo sapi
4	747	92.8	840	89 AF124993	AF124993 Homo sapi
5	729	90.6	788	89 AF231705	AF231705 Homo sapi
6	557	69.2	835	89 AF112212	AF112212 Homo sapi
7	528	65.6	780	89 AF197952	AF197952 Homo sapi
8	489	60.7	489	93 HSA249483	AJ249483 Homo sapi

9	430	53.4	601	9	A99104	A99104 Sequence 8
10	310	38.5	604	9	A99105	A99105 Sequence 9
11	263	32.7	469	9	A99103	A99103 Sequence 7
12	225	28.0	49259	78	AC090391	AC090391 Homo sapi
13	225	28.0	66266	82	AP000573	AP000573 Homo sapi
14	225	28.0	151747	82	AP001453	AP001453 Homo sapi
15	225	28.0	168475	63	AC015698	AC015698 Homo sapi
16	225	28.0	174562	60	AC005848	AC005848 Homo sapi
17	208	25.8	569	97	HS082616	HS082616 Human mRNA
18	197	24.5	2710	9	A99106	A99106 Sequence 10
19	173	21.5	157974	82	AP001380	AP001380 Homo sapi
20	145	18.0	725	88	AF110736	AF110736 Cercopithec
21	107	13.3	725	88	AF110734	AF110734 Papio ham
22	62	7.7	538	97	HS082615	HS082615 Human mRNA
23	38	4.7	798	7	AF110735	AF110735 Sus scrofa
24	38	4.7	195316	71	AC044918	AC044918 Homo sapi
25	37	4.6	195316	71	AC044918	AC044918 Homo sapi
26	36	4.5	177416	62	AC011998	AC011998 Homo sapi
27	32	4.0	675	9	A99101	A99101 Sequence 5
28	32	4.0	675	94	AF110733	AF110733 Mus muscu
29	32	4.0	688	94	AF110733	AF110733 Mus muscu
30	32	4.0	837	94	AF110732	AF110732 Rattus no
31	32	4.0	881	7	AF305564	AF305564 Bos tauru
32	32	4.0	895	94	AF124994	AF124994 Mus muscu
33	32	4.0	1869	94	AB041573	AB041573 Mus muscu
34	32	4.0	2360	94	AF208729S2	AF208729 Mus muscu
35	29	3.6	780	9	A99099	A99099 Sequence 3
36	25	3.1	25	9	A99107	A99107 Sequence 11
37	24	3.0	24	9	A99108	A99108 Sequence 12
38	24	3.0	24	9	A99109	A99109 Sequence 13
39	24	3.0	24	9	A99110	A99110 Sequence 14
40	24	3.0	24	9	A99114	A99114 Sequence 18
41	22	2.7	22	9	A99112	A99112 Sequence 16
42	22	2.7	91875	91	AL512286	AL512286 Human DNA
43	22	2.7	114411	97	HUAC002551	AC002551 Human Chr
44	22	2.7	127234	78	ALJ39155	ALJ39155 Homo sapi
45	22	2.7	186416	62	AC011226	AC011226 Homo sapi

RESULT	1
LOCUS	A99097
DEFINITION	Sequence 1 from Patent WO9909054.
ACCESSION	A99097
VERSION	A99097.1
KEYWORDS	GI:6782047
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 805)
JOURNAL	Falmane, P., Wattiez, R., Bernard, A., Hermans, C. and Knoops, B.
FEATURES	Peroxisome-associated polypeptide, nucleotide sequence encoding
source	said polypeptide and their uses in the diagnosis and/or the
CDS	treatment of lung injuries and diseases, and of oxidative
	stress-related disorders
	Patent: WO 9909054-A 1 25-FEB-1999;
	UNIV MONS HAINAUT (BE) ; FALMAGNE PAUL (BE) ; WATTIEZ RUDDY (BE) ;
	BERNARD ALFRED (BE) ; HERMANS CEDRIC (BE) ; KNOOPS BERNARD (BE) ;
	UNIV LOUVAIN (BE)
	Location/Qualifiers
	1..805
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	193...681
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAB70589.1"
	/db_xref="GI:6782048"
	/translation="MAPIKVGDAIPAVEFEGEPGNKYNLAELFKGKRGVLFGVPGF"

TPGSKTHLPQFVEQAEALAKAGVQVVAACLSVNDFAVTGEMGRAHKAAGKVRLLADPT	
GAFGKETDILLDDSLVSI FGNRLKRF SMVVDGIVKALNVEPDGTGLTCSLAPNIIS	
QL"	
BASE COUNT	163 a 200 c 276 g 166 t
ORIGIN	
Query Match	100.0%; Score 805; DB 9; Length 805;
Best Local Similarity	100.0%; Pred. No. 0;
Matches	805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 gccaggagggagtgaaagtgccgtggggcggtatggagactagctggcgtgtgcgc 60
Db	1 GCCAGGAGGGGAGTGGAAAGTGGCGGTGGCGGGTATGGGACTAGCTGGCGTGGCGCC 60
Qy	61 ctgaagacgtcagcgggctatatctctgctggggcggtggcggtcagctgcgcgcagcg 120
Db	61 CTGAGACGCTACGGGGCTATATCTCGTGGTGGGCGCGGCTCAGTCTGGCGGAGCG 120
Qy	121 gcagcaagacggtgcagtgagagagtgagctgtgctggggcggtggcggtcgcgcagcaga 180
Db	121 GCAGCAAGACGGTGCAGTGAAGGAGAGTGGCGGCTCTGGCGGGTCCGCGAGTTTCAGCAGA 180
Qy	181 gccgctgcagcgaatgcccccaatcaagggtgggagatgccatccacagcagtcgaggtgttt 240
Db	181 GCCGCTGCAGCCATGGCCCCAATCAAGGTGGGAGATGCCATCCACGACGTCGAGGTGTTT 240
Qy	241 gaaggggagccagggggaaagtgaaacctggcagagctgttcaagggaagaaggggtgtg 300
Db	241 GAAGGGAGCCAGGGGAAACAAAGGTGAACCTGGCAGAGCTGTTCAGGGGAAAGAGGTGTG 300
Qy	301 ctgtttgagtttccctggggccttccaccttcacccctgagtttcccaagacacacctgcagggttt 360
Db	301 CTGTTTGGAGTTTCTCTGGGGCTTCCACCTCGAGTGTTCACAGACACACCTGCCAGGGTTT 360
Qy	361 gtggagcaggtgaggtctctgaaggccaaagggagtcaggtggcggtgtgtgtgtgtgtgt 420
Db	361 GTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGAGTCCAGGTGGTGGCGCTGTCTGAGTGT 420
Qy	421 aatgatgcctttgactggcgagtgaggcgagcccaagggcggaagggcagaggttcgg 480
Db	421 AATGATGCTTTGTGACTGGCGAGTGGGGCGCGAGCCCAAGGCGGAAGCAAGGTTTCGG 480
Qy	481 ctctggctgacccactggggcctttgggaagagagagagacttactagatgattcgc 540
Db	481 CTCCTGGCTGATCCCACTGGGGCTTTGGGAAGGAGAGACACTATTACTAGATGATTCG 540
Qy	541 ctggtgtccatcttgggaatcgacgtctcaagaggttctccatggtgtgtgtgtgtgtgtgt 600
Db	541 CTGTTGTCCATCTTTGGGAATCGACGCTCTCAAGAGGTTCCTCCATGTTGTGTACAGGATGCC 600
Qy	601 atagtgaaggccctgaatgtggaaccagatggcacagggcctcactgcagcctgcgcaccc 660
Db	601 ATAGTGAAGCCCTGAATGTGGAAACCAAGATGGCACAGGCTTCACCTGCAGCCTGGCACCC 660
Qy	661 aatatcatctcagcgtctgagggcctggcgagattacttctccacccctccctctat 720
Db	661 AATATCATCTCAGCGCTGAGGCGCTGGCGCAGATTAATCTCTCCACCCCTCCCTATCT 720
Qy	721 cactgtccagcctgtgctggggcctcgaattggaatgtgtggcagatttctgcaata 780
Db	721 CACCTGCCAGCCCTGTGCTGGGGCTGCAATTGGAATGTTGGCCAGATTTCTGCAATA 780
Qy	781 aacactgt 805
Db	781 AACACTGT 805

RESULT	2
LOCUS	AF110731
DEFINITION	Homo sapiens antioxidant enzyme B166 mRNA, complete cds.
ACCESSION	AF110731
LOCUS	AF110731
DEFINITION	Homo sapiens antioxidant enzyme B166 mRNA, complete cds.
ACCESSION	AF110731

181 gccgcctgcagccatgccccccatcaggctgggagatgccatcccagcagtgaggctgtt 240

BASE COUNT
OPTGNBASE COUNT
OPTGN

QY 541 ctgggtccatcttgggaatcgacgtctcaagaggttctccatggtggtacagatggc 600
Db 574 CTGGGTCCATCTTTGGGAATCGACGCTCTCAAGAGGTTCTCCATGGTGTACAGATGGC 633
QY 601 atagtgaaggccctgaatgtgaaccagatgacacagcctccacctgcagctggcacc 660
Db 634 ATAGTGAAGGCCCTGAATGTGGAACAGATGGACAGGCCCTCACTGAGCTGGCACC 693
QY 661 aatatactcacagctctgagccctggccagattacttctccaccctccctatct 720
Db 694 AATATCATCTACAGCTCTGAGGCCCTGGCCAGATTAATCTCTCCAGCCCTCCCTATCT 753
QY 721 caactgccagccctgtgtggggccctccaattggaatgttgccagatttctgaata 780
Db 754 CACCTGCCAGCCGTGTGTGGGGCCCTGCAATTGGAATGTTGGCCAGATTTCTGCATA 813
QY 781 aacactgtggtttggc 798
Db 814 AACACTGTGTTGGG 831
RESULT 5
AF231705 788 bp mRNA PRI 27-JUN-2000
LOCUS Homo sapiens Alu co-repressor 1 (ACR1) mRNA, complete cds.
DEFINITION AF231705
ACCESSION AF231705
VERSION AF231705.1 GI:8745393
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 788)
AUTHORS Kropotov,A., Sedova,V., Ivanov,V., Sazeeva,N., Tomilin,A.,
Krutilina,R., Oei,S.L., Griesenbeck,J., Buchlow,G. and Tomilin,N.
A novel human DNA-binding protein with sequence similarity to a
subfamily of redox proteins which is able to repress
RNA-polymerase-III-driven transcription of the Alu-family
retroposons in vitro
JOURNAL Eur. J. Biochem. 260 (2), 336-346 (1999)
MEDLINE 99195471
PUBMED 10095767
REFERENCE 2 (bases 1 to 788)
AUTHORS Kropotov,A.V. and Tomilin,N.V.
Direct Submission
TITLE Submitted (04-FEB-2000) Chromosome Stability, Institute of Cytology
JOURNAL RAS, Tikhoretskii Av.4, St.Petersburg 194064, Russia
FEATURES
Location/Qualifiers
1..788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..788
/gene="ACR1"
20..664
/gene="ACR1"
/codon_start=1
/product="Alu co-repressor 1"
/protein_id="AAF7889.1"
/db_xref="GI:8745394"
/translation="MGLAGVCLRRSAGYILVGGAGGQSAAAARRCSGEWASGGVR
SFSRAAAMAPIKVGDPAIVAVEFVEPFGNKLAEFLKGVLFVGPAGFTPGSK
THLPFGVEQALKAQGVQVCLSVNDVFTGEWRAHKAEGKVRLLADPTGAFGKE
TDLLDLSLVSIFCNRRLKRFSSVMVQDGIKALNVEPDGTGLTCSLAPNIISQL"
BASE COUNT 153 a 201 c 268 g 166 t
ORIGIN
Query Match 90.6%; Score 729; DB 89; Length 788;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 779; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 agtggcgtggggcggtatgggaactagctggcggtgtgcccctgagacgctcagcgggc 78
Db 2 AGTGGCGTGGGGCGGGTATGGGACTAGCTGGCGTGTGGCCCTGAGAGCTCAGCGGC 61
QY 79 tatatactctcgttggggcggtcagctcgtgcggcagcgacgaagacggtcagct 138
Db 62 TATATACTCTCGTGGGGCGGGTCACTGCGGCGACGGCAGCAACACGGTGCAGT 121
QY 139 gaagagagtggtggcggttcgggggtccgcagtttccagcagagccgctgcagccatggcc 198
Db 122 GAAGAGAGTGGGGCTCTGGGGGGTCCGCAGTTTCAAGCAGAGCGCGCTGCAGCATGGCC 181
QY 199 ccaatcaaggtgggagatgccccagcggtggaggtgttgaagggaagcagcagaac 258
Db 182 CCAATCAAGGTGGGAGATGCCATCCACAGAGTGGAGGTGTTGAAGGGGAGCCAGGAAC 241
QY 259 aagtgaaacctggcagagctgtcaagggtcaagggtgtgctgtgttggagttccctggg 318
Db 242 AAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGGGTGTGCTGTTGGAGTTCCTGGG 301
QY 319 gccttcacccctggatgttccaagacacacctgcaggggttggaggcggtcgaagct 378
Db 302 GCCITCACCCCTGGATGTTCCAAGACACACCTGCACGGTGTGGAGCAGGCTGAGGCT 361
QY 379 ctgaaggccagggagtcaccaggtgtggtcgtctcgtgttaatgatgcctttgtgact 438
Db 362 CTGAAGGCCAAGGGAGTCCAGGTGGTGGCCTGCTCTGAGTGTAAATGATGCCCTTTGTGACT 421
QY 439 ggcagtggtggcgccagcccaagcggaagggttcggcctcctggctggtatccccact 498
Db 422 GCGAGTGGGGCGGAGCCCAAGCGGAGCAAGGTTCGGCTCCTGGCTGATCCCACT 481
QY 499 gggcctttgggaaggagacagacttatactagatgattcgtggtggttcacatttggg 558
Db 482 GGGCGCTTTGGGAAGAGACAGACTTATTACTAGATGATTCCTGGTGTCCATCTTTGGG 541
QY 559 aatcagctctcaagagtttccatggtgtcacaggtggtacaggtgacaggtgacaggtgact 618
Db 542 AATCGACGCTCAAGAGGTTCCTCATGGTGGTACAGGATGGCATAGTAGAGCCCTGAAT 601
QY 619 gtggaaccagatggcacagcctcacctgcagctggcaccacataatcattcacaagctc 678
Db 602 GTGGAACAGATGGCAGCGCTCACCCTGAGCCTGGCAGCAATATATCTTCACAGCTC 661
QY 679 tgaggccctgggcccagattacttccaccctccctcctcctcctcctcctcctcctcct 738
Db 662 TGAGGCCCTGGGCGAGATTAATCTCTCCACCCTCCCTATCTACCTGCCCCAGCCCTGTG 721
QY 739 ctgggcccctgcaattggaatgtggccagatttctgcaataaacaacttgggttggg 798
Db 722 CTGGGGCCCTGCANTTGGAAATGTTGCCAGATTTCTTGCAATAAACACTTGTGTTGCGG 781
RESULT 6
AF112212 835 bp mRNA PRI 13-DEC-1999
LOCUS Homo sapiens putative peroxisomal antioxidant enzyme mRNA, complete
DEFINITION cds.
ACCESSION AF112212
VERSION AF112212.1 GI:6563211
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 835)
AUTHORS Ren,S., Shi,J., Huang,C., Jiang,C., Li,Y., Zhou,J., Yu,Y., Xu,S.,
Wang,Y., Fu,G., Chen,Z. and Han,Z.
A novel gene expressed in human adrenal gland
TITLE Unpublished
JOURNAL 2 (bases 1 to 835)
REFERENCE Ren,S., Shi,J., Huang,C., Jiang,C., Li,Y., Zhou,J., Yu,Y., Xu,S.,
Wang,Y., Fu,G., Chen,Z. and Han,Z.

TITLE Direct Submission
JOURNAL Submitted (10-DEC-1998) Chinese National Human Genome Center at Shanghai, Guo Shoujing Rd. 351, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China

FEATURES
Source Location/Qualifiers
1. .835
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
184. .672
/note="PLP"
/codon_start=1
/evidence=not_experimental
/product="putative peroxisomal antioxidant enzyme"
/protein_id="AAF17200.1"
/db_xref="GI:6563212"
/translation="MAPIKVGDAIPAVEFECEPGKNVLAELFKGKGVLFQVPGAF
TPGCKSTHLPQVEQNEALKAGQVQVACLSYNDAPVTGEMGRAHKAEGKRVLLADPT
GARGKETDLLDDSLVSIIFGNRLKRFMSVMVDGIVKALNVEPDGTGLTCLAPNIIS
QL"

BASE COUNT 198 a 199 c 272 g 166 t
ORIGIN

Query Match 69.2%; Score 557; DB 89; Length 835;
Best Local Similarity 100.0%; Pred. No. 5.6e-311;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 aagacgtgcagtgaggagtggtgctgctggcggtggtccagcttcacagagcgc 185
|||||
Db 117 AAGACGCTGCAGTGAAGGAGTGGCGCTGGCGGGTCCGAGTTTCAGCAGAGCCGC 176
|||||

QY 186 tcagccatgcccacatcaagtgaggagatgccatccacagcagtgaggtgtttgaagg 245
|||||

Db 177 TCAGCCATGGCCCCAATCAAGTGGAGATGCCATCCAGCAGTGGAGTGTTCGAAGG 236
|||||

QY 246 ggagccagggaacaagtgaaacctggcagcagctgttcaaggcacaaggggtgctgtt 305
|||||

Db 237 GGAGCCAGGGAACAAGTGAACCTGGCAGAGCTGTCAAGGGCAAGAGGGTGTCTGT 296
|||||

QY 306 tggagttcctggggccttaccctggatgttcccaagacacacctgccaggtttgtgga 365
|||||

Db 297 TGGAGTCTCTGGGGCTTCCACCCCTGGATGTTCCAAAGACACACACCTGCCAGGTTTGTGGA 356
|||||

QY 366 gcaggtgaggtctgaagccaaaggagtgatcagctgtgacctgtctgagtgtaataga 425
|||||

Db 357 GCAGGTGAGGCTCGAAGGCCAAGGGATCCAGGTGGTGGCTGTCTGAGTGTATGA 416
|||||

QY 426 tgctttgtgactggcagtggtggccgagcccaagcggaaggcgaaggttcggtcct 485
|||||

Db 417 TGCCCTTGTGACTGGCGAGTGGGGCCGAGCCACAAAGCGGAAGGTTCCGGCTCT 476
|||||

QY 486 ggctgatccactgggcttggaaagagacagacttattactagatgattcggtggt 545
|||||

Db 477 GGTGATCCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATTCGCTGGT 536
|||||

QY 546 gtccatctttggaaatgacgtctcaagaggttctccatggtgtgacagatgacatagt 605
|||||

Db 537 GTCCATCTTTGGGAATCAGCTCTCAAGAGGTTCCTCATGGTGTACAGATGTCATAGT 596
|||||

QY 606 gaagccctgaatggaacagatggcacaaggcctcacctgcagcctggcaccacaatat 665
|||||

Db 597 GAAGGCCCTGATGTGGAAACAGATGGCACAGGCCTCACCTGCAGCCTGGCACCAATAT 656
|||||

QY 666 catctcacagctctgag 682
|||||

Db 657 CATCTCACAGCTCTGAG 673
|||||

RESULT 7
AF197952 780 bp mRNA PRI 29-FEB-2000
LOCUS Homo sapiens thioredoxin peroxidase PMP20 mRNA, complete cds.

ACCESSION AF197952
VERSION AF197952.1 GI:6166492
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 780)
AUTHORS Zhou, Y., Kok, K.H., Chun, A.C., Wong, C.M., Wu, H.W., Lin, M.C., Fung, P.C., Kung, H. and Jin, D.Y.
TITLE Mouse peroxiredoxin V is a thioredoxin peroxidase that inhibits p53-induced apoptosis
JOURNAL Biochem. Biophys. Res. Commun. 268 (3), 921-927 (2000)
MEDLINE 20145535
REFERENCE 2 (bases 1 to 780)
AUTHORS Jin, D.-Y.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1999) Institute of Molecular Biology, The University of Hong Kong, 8 Sassoon Road, Hong Kong
FEATURES
Location/Qualifiers
1. .780
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
110. .754
/note="type V peroxiredoxin"
/codon_start=1
/product="thioredoxin peroxidase PMP20"
/protein_id="AAF04856.1"
/db_xref="GI:6166493"
/translation="MGLAGVLCALRRSAGVILVGGAGGQSAARRCSEGEWASGGVRSFSRAAAMAPIKVGDAIPAVEFECEPGKNVLAELFKGKGVLFQVPGAFTPGCKSTHLPQVEQNEALKAGQVQVACLSYNDAPVTGEMGRAHKAEGKRVLLADPTGARGKETDLLDDSLVSIIFGNRLKRFMSVMVDGIVKALNVEPDGTGLTCLAPNIISQL"

BASE COUNT 147 a 203 c 278 g 152 t
ORIGIN

Query Match 65.6%; Score 528; DB 89; Length 780;
Best Local Similarity 99.8%; Pred. No. 3.5e-294;
Matches 678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gccaggaggcggagtggaagtgccgtggggcggtggtatgggactagctgcgtggtcgc 60
|||||

Db 74 GCCAGGAGGCGGAGTGAAGTGGCGGTGGGGGGTATGGGACTAGCTGGCGTGGCGCC 133
|||||

QY 61 ctgagacgtccagcgggtctatactctgtgtggggcggtcagtcgtcggcagcg 120
|||||

Db 134 CTGAGACGCTCAGCGGGCTATATACCTCGTGGTGGGGCGGGGTCAGTCTCGGGCAGCG 193
|||||

QY 121 gcagcaagacggtgcagtgaaaggagtggtggcgtctggcggttcgcagtttcagcaga 180
|||||

Db 194 GCAGCAAGACGGTGCAGTGAAGGAGAGTGGGGCTCTGGCGGGGTCGCGAGTTCAGCAGA 253
|||||

QY 181 gccgtgcagccatggccccaatcaagtggtgagatgccatcccgagcagtgaggtgttt 240
|||||

Db 254 GCGCTGCAGCCATGCCCCCAATCAAGGTGGGAGATGCCATCCAGCAGTGGAGGTGTT 313
|||||

QY 241 gaagggagccaggaacaagtggaacctggcagagctgttcaaggggcaagaggtgtg 300
|||||

Db 314 GAAGGGACCCAGGGAACAGGTGAACCTGGCAGAGCTGTTCAGGGCAAGAGGTGTG 373
|||||

QY 301 ctgtttggagttcctggggccttcaccccttgatgttccaaagacacacctgcaggtttt 360
|||||

Db 374 CTGTTTGGAGTTCCTGGGGCTTCCCTCGATGTTCCAGACACACCTGCCAGGTTT 433
|||||

QY 361 gtggagcagctgagctctgaagcccaagggagtcacgggtgtgtgctgtctgagtggt 420
|||||

Db 434 GTGGAGCAGGCTGAGGCTCTGAAGGCCCAAGGAGTCCAGGTGGTGGCTGTCTGAGTGT 493
|||||

QY 421 aatgatgcctttgtgactggcgagtggtggggcgagcggcacaagggcggaaggttcgg 480
|||||


```
Db 494 AATGATGCTTTGTACTGGCGAGTGGGCGGAGCCACAAAGGCGGAAGGCTTCGG 553
QY 481 ctctggctgacccactgggcccctttgggaagagacagacttattactagatgattcg 540
Db 554 CTCCTGGCTGATCCACATGGGCGCTTTTGGGAAGGAGACAGACTTATTACTAGATGATTCG 613
QY 541 ctgggtccatcttgggaatcagctctcaagaggttctccatgggtggtacagatggc 600
Db 614 CTGGGTCTCATCTTTGGGAATCGAGCTCTCAAGAGTTCTCCATGGGTGACAGATGGC 673
QY 601 atagtgaaggccctgaatgtggaacagatggcacagccctcactgcagctggcaacc 660
Db 674 ATAGTGAAGGCCCTGAATGTGGAACAGATGSCACAGGCCCTCACCTGACGCTGGCAGCC 733
QY 661 aatateatcacagctctga 681
Db 734 AATATCATCTCACAGCTCTGA 754

RESULT 8
HSA249483 489 bp mRNA PRI 02-DEC-1999
LOCUS Homo sapiens mRNA for human thiol peroxidase homologous protein
DEFINITION (TPX VI gene).
ACCESSION AJ249483
VERSION AJ249483.1 GI:6523288
KEYWORDS human thiol peroxidase homologous protein; TPX VI gene.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Kim, I. H. and Jeong, W.
JOURNAL A new type of human thiol peroxidase (Human TPx Type VI)
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 489)
TITLE Kim, I. H.
JOURNAL Direct Submission
FEATURES Submitted (13-SEP-1999) Kim I. H., Biochemistry, Paichai University,
source 439-6 Doma-2-Dong Seo-GU,, Taejon 302-735, REPUBLIC OF KOREA
1. .489 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HeLa Cell"
1. 489
/gene="TPX VI"
1. .489
/gene="TPX VI"
/function="peroxidase"
/codon_start=1
/evidence="experimental"
/product="human thiol peroxidase homologous protein"
/protein_id="CAB62210.1"
/db_xref="GI:6523289"
/translation="MAPIKVGDAIPAVEFEGEPGNKNLAEFKGKGVLFPGVPGAF
TPGCSKTHLPFGVEQAEALAKAGQVQVACLSYNDAFVTGEWGRAHKAEGKVRLLADPT
GAFGKETDILLDDSLVSIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNIIS
QL"
BASE COUNT 107 a 114 c 163 g 105 t
ORIGIN

Query Match 60.7%; Score 489; DB 93; Length 489;
Best Local Similarity 100.0%; Pred. No. 1.4e-271;
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 atggccccaatcaaggtggagatgccatccagacagtgaggtgtttgaagggagcca 252
Db 1 ATGGCCCCAATCAAGGTGGGAGATGCCATCCAGCAGTGAGGTGTTGAAGGGAGCCA 60
QY 253 gggacaaggtgaacctggcagactgttcaagggcaagaggggtgtgctgtttgagtt 312
|||||
```

```
Db 61 GGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGGGTGTGCTGTTGGAGTT 120
QY 313 cctggggccttcacccttgatgttccaagacacacctgccagggtttgtgagcaggct 372
Db 121 CTTGGGGCTTACCCCTGGATGTTCCAGACACACACTGCCAGGGTTTGTGGAGCAGGCT 180
QY 373 gaggtctgaagcccaagggagtcctgggtggcctgtctgagtgatgagtgatgagcttt 432
Db 181 GAGGCTCTGAAGGCCAAGGAGTCCAGGTGGTGGCCCTGTCTGAGTGTATTAATGATGCTTT 240
QY 433 gtactggcagatggggccagccacacagcggaaggaaggttcggctcctcgtgctgat 492
Db 241 GTGACTGGCAGTGGGGCCGAGCCACAGCCGAGGAGGACAGGTTTGGCTCTCGCTGAT 300
QY 493 cccactgggccccttgggaaggagacagacttattactagatgattcgctgggtgctcatt 552
Db 301 CCCACTGGGCCCTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCTGGTGCATC 360
QY 553 ttgggaatcgacgtctcaagaggttctccatgggtggtacaggtggcagatggaagcc 612
Db 361 TTTGGGAATCGACCTCTCAAGAGGTTCTCCATGGTGTGATGAGGATGATGATGAGGCC 420
QY 613 ctgaatggaacacagatggcacagcctcacctgcagcctggcaccacacatacatctca 672
Db 421 CTGAATGTGGAACCATGATGGCACAGGCTCACCTGCAGCCCTGGCACCAATATCATCTCA 480
QY 673 cagctctga 681
Db 481 CAGCTCTGA 489

RESULT 9
A99104 601 bp DNA PAT 20-SEP-2000
LOCUS Sequence 8 from Patent WO9909054.
DEFINITION A99104
ACCESSION A99104
VERSION A99104.1 GI:6782055
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 601)
TITLE Falmagne, P., Wattiez, R., Bernard, A., Hermans, C. and Knoops, B.
PEROXISOME-ASSOCIATED POLYPEPTIDE, NUCLEOTIDE SEQUENCE ENCODING
said polypeptide and their uses in the diagnosis and/or the
treatment of lung injuries and diseases, and of oxidative
stress-related disorders
JOURNAL Patent: WO 9909054-A 8 25-FEB-1999;
UNIV MONS HAINAUT (BE); FALMAGNE PAUL (BE); WATTIEZ RUDDY (BE);
BERNARD ALFRED (BE); HERMANS CEDRIC (BE); KNOOPS BERNARD (BE);
UNIV LOUVAIN (BE)
FEATURES Location/Qualifiers
source 1. .601
/organism="Homo sapiens"
/db_xref="taxon:9606"
161. .514
/notice="unnamed protein product"
/codon_start=1
/protein_id="CAB70593.1"
/db_xref="GI:6782056"
/translation="MAPIKTHLPFGVEQAEALAKAGQVQVACLSYNDAFVTGEWGRAH
KAEGKVRLLADPTGAFGKETDILLDDSLVSIFGNRLKRFMSVVDGIVKALNVEPDG
TGLTCSLAPNIISQL"
BASE COUNT 115 a 162 c 200 g 124 t
ORIGIN

Query Match 53.4%; Score 430; DB 9; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.1e-237;
Matches 430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 339 caagacacacctgccagggtttgtgagcaggtgaggtctgtaagggaggtcca 398
|||||
```

[illegible]

```

RESULT 10
A99105
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches 310; Conservative
98.5%; Score 310; DB 9; Length 604;
100.0%; Pred. No. 7e-168;
0; Mismatches 0; Indels 0; Gaps 0;

```

QY	33	gggtatgggactagctggcggtgcgcccctgagacgctcagcgggctatatactcgtcgg	92
Db	1	GGGTATGGGACTAGCTGGCGGTGCGCCCTGAGACGCTCAGCGGGCTATATACTCTGTCGG	60
QY	93	tggggccgcgcgtcagctctgcgcgcagcgcaagacaggtgcagtgaaagagagtgggc	152
Db	61	TGGGCCCGCGGTCACTCTGCGGCACGCGCAGCAAGACGCTGCAGTGAAGGAGAGTGGGC	120
QY	153	gtctggcgggggtccgcagtttcagcagagacgcgtgcagccatggccccaatcaaggtggg	212
Db	121	GTCTGGCGGGGTCCGCAGTTCAGCAGAGCGCTGCAGCCATGCCCCCAATCAAGGTGGG	180
QY	213	agatgccatccacagcagtgagaggtgtttcaaggggagccaggaacaaagtgaacctggc	272
Db	181	AGATGCCATCCACGACAGTGGAGGTGTTTGAAGGGGAGCCAGGGAAAGGTGAACCTGGC	240
QY	273	agagctgttcaaggccaagaaaggtgtgctgtttgagttcctgggggcctcaccacctgg	332
Db	241	AGAGCTGTTCAGGGCAAGAGGGTGTGCTGTTTGGAGTTCCTGGGGCCTTCACCCCTGG	300
QY	333	atgttccaag	342
Db	301	ATGTTCCTCAAG	310

RESULT	11
LOCUS	A99103
DEFINITION	Sequence 7 from Patent WO9909054.
ACCESSION	A99103
VERSION	A99103.1
KEYWORDS	GI:6782053
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 469)
AUTHORS	Falmagne,P., Wattiez,R., Bernard,A., Hermans,C. and Knoops,B.
TITLE	Peroxisome-associated polypeptide, nucleotide sequence encoding said polypeptide and their uses in the diagnosis and/or the treatment of lung injuries and diseases, and of oxidative stress-related disorders
JOURNAL	Patent: WO 9909054-A 7 25-FEB-1999; UNIV MONS HAINAUT (BE) ; FALMAGNE PAUL (BE) ; BERNARD ALFRED (BE) ; HERMANS CEDRIC (BE) ; KNOOPS BERNARD (BE) ; UNIV LOUVAIN (BE) Location/Qualifiers 1 469
FEATURES	source

```

source
1. .469
/organism="Homo sapiens"
/db_xref="taxon:9606"
161. .382
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB70592.1"
/db_xref="GI:6782054"
/translation="MAPIKVRLLADPTGAFKETDLLDLSLVIFGNRRLLKRFMSVV
QDGIVKALNVEPDGTLGTCISLAPNISQL"
88 a 132 c 150 q 99 t
BASE COUNT

```

	Query Match	32.7%	Score 263;	DB 9;	Length 469;	
	Best Local Similarity	100.0%;	Pred. No. 1.2e-140;			
	Matches 263;	Conservative	0;	Mismatches	0;	Indels
						Gaps
QY	471	caaggttcgctcctgctgatccactcctgggccccttgggaaggagacagacttatt	530			
Db	172	CAAGGTTCCGCTCCTGCTGATCCCACTGGGGCCCTTGGGAAGGAGACACACTTATTACT	231			
QY	531	agatgattcctcctgggtcccatcttttgggaatcagctctc	590			
Db	232	AGATGATTCGCTGGTGTCACATCTTTGGGAATCCAGCTCTCAAGAGGTTCTCCATGGTGGT	291			

QY 591 acaggtgcatagtgagggccctgaatgtggaaccagatggcacaggccctcacctgcag 650
 |||||
 Db 292 ACAGGATGCATAGTGAAGGCCCTGAATGTGGAACAGATGGACAGGCCCTCACCTGCAG 351
 |||||

QY 651 cctggcaccacaatactatctacagctctgagcgccctggccagattcttctctccaccc 710
 |||||
 Db 352 CTTGGCACCANTATCTACAGCTCTGAGGCCCTGGCCAGATTACTTCTCTCCACCC 411
 |||||

QY 711 ctccctatctacccgtccagcc 733
 |||||
 Db 412 CTCCCTATCTACCTGCCAGCC 434
 |||||

RESULT 12
 AC090391
 LOCUS
 DEFINITION Homo sapiens chromosome 11 clone RP11-87716 map 11, LOW-PASS
 SEQUENCE SAMPLING.
 AC090391
 AC090391.1 GI:12958027
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 11, clone RP11-87716
 Unpublished
 2 (bases 1 to 49259)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
 Barina,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,
 Camarata,J., Campopiano,A., Choepel,V., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
 Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
 Lechoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
 Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
 McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
 Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
 Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
 Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Travers,M., Travis,N., Trigglio,J., Vassiliev,H., Viel,R., Vo,A.,
 Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 COMMENT
 Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L12693
 Center clone name: 877_I_6

* NOTE: This record contains 62 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 702 801: contig of 701 bp in length
 802 1496: contig of 695 bp in length
 1497 1596: gap of 100 bp
 1597 2252: contig of 656 bp in length
 2253 2352: gap of 100 bp
 2353 3061: contig of 709 bp in length
 3062 3161: gap of 100 bp
 3162 3869: contig of 708 bp in length
 3870 3969: gap of 100 bp
 3970 4681: contig of 712 bp in length
 4682 4781: gap of 100 bp
 4782 5514: contig of 733 bp in length
 5515 5614: gap of 100 bp
 5615 6299: contig of 685 bp in length
 6300 6399: gap of 100 bp
 6400 7087: contig of 688 bp in length
 7088 7187: gap of 100 bp
 7188 7862: contig of 675 bp in length
 7863 7962: gap of 100 bp
 7963 8673: contig of 711 bp in length
 8674 8773: gap of 100 bp
 8774 9484: contig of 711 bp in length
 9485 9584: gap of 100 bp
 9585 10302: contig of 718 bp in length
 10303 10402: gap of 100 bp
 10403 11104: contig of 702 bp in length
 11105 11204: gap of 100 bp
 11205 11921: contig of 717 bp in length
 11922 12021: gap of 100 bp
 12022 12698: contig of 677 bp in length
 12699 12798: gap of 100 bp
 12799 13513: contig of 715 bp in length
 13514 13613: gap of 100 bp
 13614 14327: contig of 714 bp in length
 14328 14427: gap of 100 bp
 14428 15106: contig of 679 bp in length
 15107 15206: gap of 100 bp
 15207 15877: contig of 671 bp in length
 15878 15977: gap of 100 bp
 15978 16670: contig of 693 bp in length
 16671 16770: gap of 100 bp
 16771 17443: contig of 673 bp in length
 17444 17543: gap of 100 bp
 17544 18233: contig of 690 bp in length
 18234 18333: gap of 100 bp
 18334 19012: contig of 679 bp in length
 19013 19112: gap of 100 bp
 19113 19806: contig of 694 bp in length
 19807 19906: gap of 100 bp
 19907 20573: contig of 667 bp in length
 20574 20673: gap of 100 bp
 20674 21371: contig of 698 bp in length
 21372 21471: gap of 100 bp
 21472 22184: contig of 713 bp in length
 22185 22284: gap of 100 bp
 22285 22976: contig of 692 bp in length
 22977 23076: gap of 100 bp
 23077 23749: contig of 673 bp in length
 23750 23849: gap of 100 bp
 23850 24526: contig of 677 bp in length
 24527 24626: gap of 100 bp
 24627 25319: contig of 693 bp in length
 25320 25419: gap of 100 bp
 25420 26077: contig of 658 bp in length
 26078 26177: gap of 100 bp
 26178 26838: contig of 661 bp in length
 26839 26938: gap of 100 bp
 26939 27644: contig of 706 bp in length

* 27645 27744: gap of 100 bp in length
 * 27745 28437: contig of 693 bp in length
 * 28438 28537: gap of 100 bp
 * 28538 29246: contig of 709 bp in length
 * 29247 29346: gap of 100 bp
 * 29347 30054: contig of 708 bp in length
 * 30055 30154: gap of 100 bp
 * 30155 30842: contig of 688 bp in length
 * 30843 30942: gap of 100 bp
 * 30943 31622: contig of 680 bp in length
 * 31623 31722: gap of 100 bp
 * 31723 32382: contig of 660 bp in length
 * 32383 32482: gap of 100 bp
 * 32483 33197: contig of 715 bp in length
 * 33198 33297: gap of 100 bp
 * 33298 33994: contig of 697 bp in length
 * 33995 34094: gap of 100 bp
 * 34095 34803: contig of 709 bp in length
 * 34804 34903: gap of 100 bp
 * 34904 35600: contig of 697 bp in length
 * 35601 35700: gap of 100 bp
 * 35701 36487: contig of 787 bp in length
 * 36488 37280: contig of 693 bp in length
 * 37281 37380: gap of 100 bp
 * 37381 38079: contig of 699 bp in length
 * 38080 38179: gap of 100 bp
 * 38180 38858: contig of 679 bp in length
 * 38859 39595: gap of 100 bp
 * 39596 39611: contig of 653 bp in length
 * 39612 39711: gap of 100 bp
 * 39712 40413: contig of 702 bp in length
 * 40414 40513: gap of 100 bp
 * 40514 41220: contig of 707 bp in length
 * 41221 41320: gap of 100 bp
 * 41321 42032: contig of 712 bp in length
 * 42033 42132: gap of 100 bp
 * 42133 42848: contig of 716 bp in length
 * 42849 42948: gap of 100 bp
 * 42949 43674: contig of 726 bp in length
 * 43675 43774: gap of 100 bp
 * 43775 44495: contig of 721 bp in length
 * 44496 44595: gap of 100 bp
 * 44596 45298: contig of 703 bp in length
 * 45299 45398: gap of 100 bp
 * 45399 46085: contig of 687 bp in length
 * 46086 46185: gap of 100 bp
 * 46186 46875: contig of 690 bp in length
 * 46876 46975: gap of 100 bp
 * 46976 47652: contig of 677 bp in length
 * 47653 47752: gap of 100 bp
 * 47753 48443: contig of 691 bp in length
 * 48444 48543: gap of 100 bp
 * 48544 49259: contig of 716 bp in length.

FEATURES

Location/Qualifiers
 1. .49259
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-87716"
 /clone_lib="RPC1-11 Human Male BAC"

BASE COUNT 10101 a 10720 c 11391 g 10620 t 6427 others
 ORIGIN

Query Match 28.08; Score 225; DB 78; Length 49259;
 Best Local Similarity 100.0%; Pred. No. 1.2e-118;
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy .574 aggtctccatggttggtacaggtggcagatgaagccctgaatggtggaacagatggc 633
 Db 14599 AGGTTCTCCATGGTGTACAGGATGGCAGTGTGAGGCCCTGGAATGGAACCATGGC 14658

QY 634 acaggcctcacctgacgctggcaccacaatcatctacagctctcagccctgggcca 693
 Db 14659 ACAGGCTCACCCTGACCCCTGGCACCACCAATATCATCTCAGGCTCTGAGGCCCTGGGCCA 14718
 QY 694 gattactctccaccctccctctatctcacctgcccagccctgtgctggggccctgcaat 753
 Db 14719 GATTACTTCTCCACCCTCCCTATCTCACCTGCCAGCCCTGTCGCTGGGCCCTGCAAT 14778
 QY 754 tggaattgtggcagattttctgcaataaacaacttggtgtggtg 798
 Db 14779 TGAATGTTGGCCAGATTTCTGCAATAAACACTTGTGTTGGCG 14823

RESULT 13

AP000573 66266 bp DNA HTG 30-MAY-2000
 LOCUS Homo sapiens chromosome 11 clone XXpl-25H10 map 11q13, WORKING
 DEFINITION DRAFT SEQUENCE, 23 unordered pieces.
 ACCESSION AP000573
 VERSION AP000573.3 GI:8118781
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens DNA, clone:XXpl-25H10.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 66266)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Homo sapiens 66,266 genomic DNA of 11q13
 JOURNAL Published Only in Database (1999) In press
 REFERENCE 2 (bases 1 to 66266)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (08-OCT-1999) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
 Japan (E-mail:hattori@gsc.riken.go.jp,
 URL:http://hqp.gsc.riken.go.jp/, Tel:81-42-778-9923,
 Fax:81-42-778-9924)
 COMMENT On May 31, 2000 this sequence version replaced gi:6997465.
 ----- Genome Center
 Center: RIKEN Genomic Sciences Center(GSC)
 Center code: RIKEN
 Web site: http://hqp.gsc.riken.go.jp/
 Contact: hattori@gsc.riken.go.jp
 ----- Project Information
 Center project name: Humdrift11
 Center clone name: XXpl-25H10
 ----- Summary Statistics
 Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 56811 bases at least Q40
 Consensus quality: 60215 bases at least Q30
 Consensus quality: 62540 bases at least Q20
 Insert size: 64066; sum-of-contigs
 Quality coverage: 4.15x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 23 contigs. The true order of the pieces is not known and the order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1	6017 contig of	6017 bp in length
6118	11749 contig of	5632 bp in length
11850	16292 contig of	4443 bp in length
16393	19696 contig of	3304 bp in length
19797	23376 contig of	3580 bp in length
23477	27638 contig of	4162 bp in length

27739 31508 contig of 3770 bp in length
31609 35096 contig of 3488 bp in length
35197 39266 contig of 4070 bp in length
39367 41973 contig of 2607 bp in length
42074 43566 contig of 1493 bp in length
43667 46358 contig of 2692 bp in length
46459 49035 contig of 2577 bp in length
49136 50654 contig of 1519 bp in length
50755 53342 contig of 2588 bp in length
53443 56951 contig of 1792 bp in length
55335 56951 contig of 1617 bp in length
57052 59082 contig of 2031 bp in length
59183 60880 contig of 1698 bp in length
60981 62607 contig of 1627 bp in length
62708 63786 contig of 1079 bp in length
63887 64901 contig of 1015 bp in length
65002 66266 contig of 1265 bp in length

Sequence updated (04-Nov-1999)
Sequence updated (26-May-2000).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
6018 6117: contig of 6017 bp in length
6118 11749: contig of 5632 bp in length
11750 11849: gap of 100 bp
11850 16292: contig of 4443 bp in length
16293 16392: gap of 100 bp
16393 19696: contig of 3304 bp in length
19697 19796: gap of 100 bp
19797 23376: contig of 3580 bp in length
23377 23476: gap of 100 bp
23477 27638: contig of 4162 bp in length
27639 27738: gap of 100 bp
27739 31508: contig of 3770 bp in length
31509 31608: gap of 100 bp
31609 35096: contig of 3488 bp in length
35097 35196: gap of 100 bp
35197 39266: contig of 4070 bp in length
39267 39366: gap of 100 bp
39367 41973: contig of 2607 bp in length
41974 42073: gap of 100 bp
42074 43566: contig of 1493 bp in length
43567 43666: gap of 100 bp
43667 46358: contig of 2692 bp in length
46359 46458: gap of 100 bp
46459 49035: contig of 2577 bp in length
49036 49135: gap of 100 bp
49136 50654: contig of 1519 bp in length
50655 50754: gap of 100 bp
50755 53342: contig of 2588 bp in length
53343 53442: gap of 100 bp
53443 55234: contig of 1792 bp in length
55235 55334: gap of 100 bp
55335 56951: contig of 1617 bp in length
56952 57051: gap of 100 bp
57052 59082: contig of 2031 bp in length
59083 59182: gap of 100 bp
59183 60880: contig of 1698 bp in length
60881 60980: gap of 100 bp
60981 62607: contig of 1627 bp in length
62608 62707: gap of 100 bp
62708 63786: contig of 1079 bp in length
63787 63886: gap of 100 bp
63887 64901: contig of 1015 bp in length
64902 65001: gap of 100 bp
65002 66266: contig of 1265 bp in length.

Location/Qualifiers

source
1. .56266
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/clone="XXp1-25H10"
1. .6017
/note="assembly_fragment"
6118. .11749
/note="assembly_fragment"
11850. .16292
/note="assembly_fragment"
16393. .19696
/note="assembly_fragment"
19797. .23376
/note="assembly_fragment"
23477. .27638
/note="assembly_fragment"
27739. .31508
/note="assembly_fragment"
31609. .35096
/note="assembly_fragment"
35197. .39266
/note="assembly_fragment"
39367. .41973
/note="assembly_fragment"
42074. .43566
/note="assembly_fragment"
43667. .46358
/note="assembly_fragment"
46459. .49035
/note="assembly_fragment"
49136. .50654
/note="assembly_fragment"
50755. .53342
/note="assembly_fragment"
53443. .55234
/note="assembly_fragment"
55335. .56951
/note="assembly_fragment"
57052. .59082
/note="assembly_fragment"
59183. .60880
/note="assembly_fragment"
60981. .62607
/note="assembly_fragment"
62708. .63786
/note="assembly_fragment"
63887. .64901
/note="assembly_fragment"
65002. .66266
/note="assembly_fragment"
BASE COUNT 13932 a 17688 c 18127 g 14319 t 2200 others
ORIGIN
Query Match 28.0%; Score 225; DB 82; Length 66266;
Best Local Similarity 100.0%; Pred. No. 1.2e-118; Indels 0; Gaps 0;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 574 aggtctccatggtgacagatgacatggaagccctgaatgtggaaccagatgac 633
|||||
Db 13048 AGGTTCTCCATGGTGTACAGGATGCGATGAGGCGCTGAGTGTGACACGATGCG 13107
QY 634 acaggctcacctgcagcctggcaccacaataatcatcatcagctctgagccctggcca 593
|||||
Db 13108 ACAGGCTTCACCTGCAGCTGGCACCAATATCATCATCAGCTCTGAGGCGCTGGGCA 13167
QY 694 gattactctcacccctccctctatctaccctgcccagccctgctgctgctgcaat 753
|||||
Db 13168 GATTACTTCTCCACCCCTCCCTATCTACCTGCCAGCCCTGTGCTGGGCCCTGCAAT 13227
QY 754 tggaaatgtggccagatttctgcaataaacacactgtggttggcgg 798

Db 13228 TGGATGTGCCAGATTCTGCAATAACACTTGTGTTGCGG 13272
 |||

RESULT 14

AP001453

LOCUS

DEFINITION

AP001453 Homo sapiens chromosome 11 clone RP11-783K16 map 11q13, WORKING
 DRAFT SEQUENCE, 37 unordered pieces.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (15-MAR-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
 (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
 Tel: 81-42-778-9923, Fax: 81-42-778-9924)
 On Sep 15, 2000 this sequence version replaced gi:8117327.

COMMENT

----- Genome Center

Center: RIKEN Genomic Sciences Center (GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

----- Project Information

Center project name: Humdraft11

Center clone name: RP11-783K16

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 136959 bases at least Q40

Consensus quality: 143151 bases at least Q30

Consensus quality: 146210 bases at least Q20

Insert size: 148147; sum-of-contigs

Quality coverage: 8.32x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
 37 contigs. The true order of the pieces is not known and their
 order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved

1
 9802 contig of 8902 bp in length
 9003 18232 contig of 9230 bp in length
 18333 28659 contig of 10327 bp in length
 28760 37331 contig of 8572 bp in length
 37432 45066 contig of 7635 bp in length
 45167 51737 contig of 6571 bp in length
 51838 57917 contig of 6080 bp in length
 58018 62697 contig of 4159 bp in length
 62798 66342 contig of 3545 bp in length
 66443 69246 contig of 2804 bp in length
 69347 72173 contig of 2827 bp in length
 72174 72773 contig of 100 bp
 72774 76432 contig of 4159 bp in length
 76433 76532 contig of 100 bp
 76533 80848 contig of 4316 bp in length
 80849 80948 contig of 100 bp
 80949 84529 contig of 3581 bp in length
 84530 84629 contig of 100 bp
 84630 88588 contig of 3959 bp in length
 88589 88688 contig of 100 bp
 88689 94224 contig of 5536 bp in length
 94225 94324 contig of 100 bp
 94325 98727 contig of 4403 bp in length
 98728 98827 contig of 100 bp
 98828 102161 contig of 3333 bp in length
 102161 102260 contig of 100 bp
 102261 107715 contig of 5455 bp in length
 107716 107815 contig of 100 bp
 107816 111480 contig of 3665 bp in length
 111481 111580 contig of 100 bp
 111581 114035 contig of 2455 bp in length
 114036 114135 contig of 100 bp
 114136 117720 contig of 3585 bp in length
 117721 117820 contig of 100 bp
 117821 121176 contig of 3356 bp in length

98828 102160 contig of 3333 bp in length
 102261 107715 contig of 5455 bp in length
 107716 111480 contig of 3665 bp in length
 111481 114035 contig of 2455 bp in length
 114036 117720 contig of 3585 bp in length
 117821 121176 contig of 3356 bp in length
 121277 123962 contig of 2686 bp in length
 124063 128424 contig of 4362 bp in length
 128525 131280 contig of 2756 bp in length
 131381 133764 contig of 2384 bp in length
 133865 137166 contig of 3302 bp in length
 137267 139968 contig of 2702 bp in length
 140069 141380 contig of 1312 bp in length
 141481 142715 contig of 1235 bp in length
 142816 144247 contig of 1432 bp in length
 144348 146022 contig of 1675 bp in length
 146123 147416 contig of 1294 bp in length
 147517 148926 contig of 1294 bp in length
 149027 150172 contig of 1410 bp in length
 150273 151747 contig of 1475 bp in length.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 37 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 8903 9002: contig of 8902 bp in length
 9003 18232: gap of 100 bp
 18233 18332: contig of 9230 bp in length
 18333 28659: contig of 10327 bp in length
 28660 28759: gap of 100 bp
 28760 37331: contig of 8572 bp in length
 37332 37431: gap of 100 bp
 37432 45066: contig of 7635 bp in length
 45067 45166: gap of 100 bp
 45167 51737: contig of 6571 bp in length
 51738 51837: gap of 100 bp
 51838 57917: contig of 6080 bp in length
 57918 58017: gap of 100 bp
 58018 62697: contig of 4159 bp in length
 62698 62797: gap of 100 bp
 62798 66342: contig of 3545 bp in length
 66343 66442: gap of 100 bp
 66443 69246: contig of 2804 bp in length
 69247 69346: gap of 100 bp
 69347 72173: contig of 2827 bp in length
 72174 72773: gap of 100 bp
 72774 76432: contig of 4159 bp in length
 76433 76532: gap of 100 bp
 76533 80848: contig of 4316 bp in length
 80849 80948: gap of 100 bp
 80949 84529: contig of 3581 bp in length
 84530 84629: gap of 100 bp
 84630 88588: contig of 3959 bp in length
 88589 88688: gap of 100 bp
 88689 94224: contig of 5536 bp in length
 94225 94324: gap of 100 bp
 94325 98727: contig of 4403 bp in length
 98728 98827: gap of 100 bp
 98828 102160: contig of 3333 bp in length
 102161 102260: gap of 100 bp
 102261 107715: contig of 5455 bp in length
 107716 107815: gap of 100 bp
 107816 111480: contig of 3665 bp in length
 111481 111580: gap of 100 bp
 111581 114035: contig of 2455 bp in length
 114036 114135: gap of 100 bp
 114136 117720: contig of 3585 bp in length
 117721 117820: gap of 100 bp
 117821 121176: contig of 3356 bp in length

Search completed: October 22, 2001, 13:44:51
Job time: 5461 sec

* as soon as it is available and the accession number will
* be preserved.

1 40714: contig of 40714 bp in length
40715 40814: gap of 100 bp
40815 42233: contig of 1419 bp in length
42234 42333: gap of 100 bp
42334 47280: contig of 4947 bp in length
47281 47380: gap of 100 bp
47381 53604: contig of 6224 bp in length
53605 53704: gap of 100 bp
53705 61923: contig of 8225 bp in length
61930 62023: gap of 100 bp
62030 109187: contig of 47158 bp in length
109188 109287: gap of 100 bp
109288 131160: contig of 21873 bp in length
131161 131260: gap of 100 bp
131261 150366: contig of 19106 bp in length
150367 150466: gap of 100 bp
150467 168475: contig of 18009 bp in length.

FEATURES

Source

1. 168475
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="Rp11-147G6"
/clone_lib="RPCI-11 Human Male BAC"
1. 40714
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
40815..42233
/note="assembly_fragment"
42334..47280
/note="assembly_fragment"
47381..53604
/note="assembly_fragment"
53705..61929
/note="assembly_fragment"
62030..109187
/note="assembly_fragment"
109288..131160
/note="assembly_fragment"
131261..150366
/note="assembly_fragment"
150467..168475
/note="assembly_fragment"
clone_end:17
vector_side:right"

BASE COUNT 40113 a 42972 c 43226 g 41363 t 801 others
ORIGIN

Query Match 28.08; Score 225; DB 63; Length 168475;
Best Local Similarity 100.08; Pred. No. 1.2e-118;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 574 aggttccatggtgtacagatggcatagtgaagccctgaatgtggaaccagatggc 633
|||||
Db 165891 AGGTTCTCATGTTGTTACAGATGGCATAGTGAAGGCCCTGGAATGTGAACACAGATGGC 165832
QY 634 acagccctcaactgcagcctggcaccacaatatcatcagctctgagccctgggcca 693
|||||
Db 165831 ACAGCCCTCACCTGCAGCCTGGCACCACCAATATCATCTCACAGCTCTGAGGCCCTGGCCA 165772
QY 694 gattacttcctcccccctccctctatctcactgcagccctgtgtgtggtttg 753
|||||
Db 165771 GATTACTTCCTCCACCCCTCCCTATCTACCTGCCACGCCCTGTGTGGGCCCTGCAAT 165712
QY 754 tggatgttggccagatttcttgcataaacaactgtgtgtttg 798
|||||
Db 165711 TGGATGTGGCCAGATTTCTTGAATAAACACATTGTGTTTGGG 165667

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2001, 12:12:15 ; Search time 1291.82 seconds
(without alignments)
5890.563 Million cell updates/sec

Title: US-09-486-167A-1

Perfect score: 805

Sequence: 1 gccaggagcgagtggaag.....ttgtgttgcgaaaaaaa 805

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 10228115 seqs, 4726426750 residues

Word size: 0

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*


```

|||||
Db 722 AATATCATCTCACAGCTGAGGCCTGGCCAGATTACTTCTCCACCCCTCCCTATCT 781
Qy 721 cacctggccagccctgtgctgggcccctgcaattggaattgttggccaga 769
|||||
Db 782 CACCTGCCAGCCCTGCTGCTGGGCCCTTGCATTTGGAATGTTGCCAGA 830
|||||

RESULT 2
AV650400 677 bp mRNA EST 07-SEP-2000
LOCUS AV650400 GLC Homo sapiens cDNA clone GLCED08 3', mRNA sequence.
DEFINITION AV650400
ACCESSION AV650400
VERSION AV650400.1 GI:9871414
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 677)
AUTHORS Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H.,
Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J.,
Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA clone
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
Location/Qualifiers
1..677
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCED08"
/clone_lib="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT 136 a 156 c 248 g 132 t 3 others
ORIGIN

Query Match 79.6%; Score 641; DB 32; Length 677;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gccaggaggcggagtggaagtggcgtggggcggtatgactagtcgctgtgcgc 60
Db 8 gccaggaggcggagtggaagtggcgtggggcggtatgactagtcgctgtgcgc 67
|||||
Qy 61 ctgagacgtcagcgggctatctactcgtcgtggggcggtcgtcgtcgtgcgcagc 120
Db 68 CTGAGACGCTCAGCGGGCTATATCTCGTGGTGGGCGCGGTCTAGTCTGGGCAGCG 127
|||||
Qy 121 qcagcaagacggtgcaatgagagagtggtgctcgtggcggttcgcagtttcagcaga 180
Db 128 GCACGAAGACGGGTGACGTGAAGGAGAGTGGGGCGTCTGGCGGGGTCCGACGTTTCAGCAGA 187
|||||
Qy 181 gcgcgtcagccatggcccaatcaagtggagatgccatccacagcagtgagggtttt 240
Db 188 GCCGCTCAGCCATGGCCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTT 247
|||||
Qy 241 gaaggggagccagggacaaggtgaacctggcagagctgttcaagggcaagaaggggtg 300
Db 248 GAAGGGGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAAGGGGTGTG 307

```

```

Qy 301 ctgtttgagttctcctggggccttcacccctgatttccaagacacacacctgccaggttt 360
Db 308 CTGTTTGAGTCTCTGGGGCCTTCACCCCTGGATGTTCCAGACACACCTGCCAGGGTTT 367
|||||
Qy 361 gtgagcagggctgaggtctctgaaggccaagggagtcacaggtggtgctgtctgagtt 420
Db 368 GTGAGACAGGCTGAGGCTCTGAAGGCCAAGGAGTCCAGGTGGCTGCTGAGTGT 427
|||||
Qy 421 aatgatgccttctgactggcggagtgggccgagcccaagggcgaaggttcgg 480
Db 428 AATGATGCTTGTGACTGGCGAGTGGGGCCGAGCCCAAGGCCGAAGGTTCGG 487
|||||
Qy 481 ctctgggtgatcccaactggggcctttgggaaggagacagacttattactagatgatcg 540
Db 488 CTCCTGGCTGATCCCACTGGGGCCTTTTGGGAAGGAGACAGACTTATTACTAGATGTCG 547
|||||
Qy 541 ctggtgtccatttttgggaatcgacgtctcaagggttctccatgtggtacagagatggc 600
Db 548 CTGGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTTCTCCATGGTGTACAGGATGCG 607
|||||
Qy 601 atagtgaagggcctggaatgtggaaccagatgacacagcct 641
Db 608 ATAGTGAAGGCCCTGAATGTGGACCATGATGACATGACATGACATGACATGACATGAC 648
|||||

RESULT 3
BE784530 790 bp mRNA EST 20-OCT-2000
LOCUS BE784530 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3876972 5',
DEFINITION mRNA sequence.
ACCESSION BE784530
VERSION BE784530.1 GI:10205815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAM9638 row: j column: 13
High quality sequence stop: 733.

FEATURES
source
Location/Qualifiers
1..790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3876972"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

BASE COUNT 161 a 193 c 273 g 163 t
ORIGIN

Query Match 78.6%; Score 633; DB 140; Length 790;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 gagtggaaagtggcgtggggcggtatggactagctggcgtgctgcgcctgagacgctc 71

```

•	1	GAGTGGAA	TGGCGTGGGGCGGTATGGACTAGCTGGCGTGTGGCCCTGTGAGACGCTC	60
Ddb				
Qy	72	agcgggtata	ctcgtcggtggggcggcggtcactctcggcacgcacacaagc	131
Ddb				
Db	61	AGCGGGCTAT	TACTCGTCGTCGGGGCGGGGTGAGTCTGCGCAGCGCAGCAAGC	120
Qy	132	gtcagatgaagagag	tggcgctgctggcggtccgcagttccagcagagccctgcgcgc	191
Ddb				
Db	121	GTGCAGTGAAG	GAGAGTGGCGTCTGCGGGGTCCGAGTTTCAGCAGAGCGCGTGCAGC	180
Qy	192	catggcccca	tcaagttggagatgccatccacgacgtggaggtgtttgaagggaagcc	251
Ddb				
Db	181	CATGCCCCCAAT	CAAGTGGGAGATGCCATCCACGAGTGGAGGTGTTGAAGGGGAGCC	240
Qy	252	agggaacaagt	gaacctggcagagctgttcaagggcaagaagggtgtgctgtttggagt	311
Ddb				
Db	241	AGGGAAACAAG	TGAACCTGGCGAGAGCTGTTCAAGGGCAAGAAGGGTGTGCTGTTGGAGT	300
Qy	312	tectggggcctt	caaccctggatgttccaagacacacctgccaggggtttgtggagcaggc	371
Ddb				
Db	301	TCTTGGGGCTT	CAACCCCTGGATGTTCCAAAGACACACTGCCAAGGGTGTGGAGCAGGC	360
Qy	372	tgaagctcgaagc	ccaagcagtcacagtgctgctgagtgatgaatgatgcctt	431
Ddb				
Db	361	TGAGGCTCTGA	AGGCCAAGGAGTCCAGGTGGTGGCCTGCTGAGTGTATGATGCCTT	420
Qy	432	tgtgactgcgagt	ggggccagcccaagcgcggaagcgaagttcggctcctggctga	491
Ddb				
Db	421	TCTGACTGCG	AGTGGGGCGAGCCACCAAGCGGAAGCTTCGGCTCCCTGGGTGA	480
Qy	492	tcccactgggcctt	ggaaagagacagacttattactagatgattcgcgtgggtgccat	551
Ddb				
Db	481	TCCCCTGGGGC	TTTGGGAAGGACAGACTTATTACTAGATGATTCGCTGGTGTCCAT	540
Qy	552	ctttgggaatc	acgactccaagaggttccatggtgacagatggcatagtgaaagcc	611
Ddb				
Db	541	CTTTGGGAAT	CACGCTCTCAAGAGGTTCTCCATGGTGGTACAGGATGGCATAGTGAAGC	600
Qy	612	cttgaatgtgga	accagatggcacagccctcac	644
Ddb				
Db	601	CTTGAATGTG	GGAACAGATGGCACAGGCCTCAC	633

RESULT	4
BG255777	
LOCUS	622 bp mRNA EST 13-FEB-2001
DEFINITION	602368257F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4476207 5', mRNA sequence.
ACCESSION	BG255777
VERSION	BG255777.1 GI:12765593
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 622) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
AUTHORS	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov
TITLE	Tissue Procurement: DCTD/DMP
JOURNAL	CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
COMMENT	DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10303 row: b column: 16 High quality sequence stop: 622. Location/Qualifiers 1..622
FEATURES	source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4476207"
/clone_lib="NIH_MGC_91"
/tissue_type="adenocarcinoma, cell line"
/lab_host="pHI10B (phage-resistant)"
/notes="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."

```

[illegible]

Db '188 GCGGCTGAGCCATGGCCCCAATCAAGGTGGGAGATCCATCCAGCAGTGGAGGTGTTT 247

Qy 241 gaagggagccagggaacaaagtgtaacccctggcagagctgttcaaggccaagaggtgtg 300

Db 248 GAAGGGAGCCAGGAAACAAAGGTGAACCTGGCCAGACGCTGTTCAAGGGCAAGAGGTGTG 307

Qy 301 ctgtttggagttctctgggggcttcccccctggatgtttccaaagacacacacctgccaggttt 360

Db 308 CTGTTTGGAGTTCCTGGGGCTTCCACCTGGATGTTTCAAGACACACACCTGCCAGGTTT 367

Qy 361 gtggagcagctgaggtctgaaggccaaggaggtccagggtgtgtcgtcgtcgtgtgt 420

Db 368 GTGAGCAGGCTGAGGCTCTGAAGGCCAAGGGAGTCCAGGTGGTGGCTCTGTAGTGT 427

Qy 421 aatgatcccttctgactggcaggtgggcccagcccaagcgaagcgaagcgaaggttcgg 480

Db 428 ATGATGCTCTTGTGACTGCGAGTGGGCGGAGCCGAGCCACAAAGCGGAAGGCTTCGG 487

Qy 481 ctctggctgacccactggggccttgggaaggagacagacttatactagatgattcg 540

Db 488 CTCTGGCTGATCCACTGGGGCTTGGGAGGAGACAGACTTATTACTAGATGATCG 547

Qy 541 ctggtgtccatcttgggaatcacgtctcaagagttctccatgtgtgtacagatgac 600

Db 548 CTGCTGCCATCTTTGGGAATCAGCTCTCAAGAGGTTCCTCCATGGTGGTACAGGATGGC 607

Qy 601 atagtgaagcc 612

Db 608 ATAGTGAAGGCC 619

RESULT 7

BE298055 604 bp mRNA EST 20-JUL-2000

LOCUS 601118369F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3028312 5',

DEFINITION mRNA sequence.

ACCESSION BE298055

VERSION BE298055.1 GI:9181641

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 604)

TITLE NIH-MGC http://mgc.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov

Plate: L1CM86 row: i column: 17

High quality sequence stop: 601.

Location/Qualifiers

1..604

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3028312"

/clone_lib="NIH_MGC_17"

/tissue_type="rhabdomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using 2AP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 123 a 158 c 194 g 129 t

ORIGIN

Query Match 75.0%; Score 604; DB 165; Length 604;

Best Local Similarity 100.0%; Pred. NO. 1.9e-301;

Matches 604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 156 tggcggggtccgcagctttccagcagagccqctgcagccatggccccaatcaagggtgggaga 215

Db 1 TGGCGGGGTCCGCAGTTTCAGCAGAGCCGCTGCAGCCATGCCCAATCAAGGTGGGAGA 60

Qy 216 tgcctccacagcagtgagggtgtttgaaggggagccaggggaacaaggtgaacctggcaga 275

Db 61 TGCCATCCAGCAGTGGAGGTGTTTGAAGGGGAGCCAGGGAACAGGTGAACCTGGCAGA 120

Qy 276 gctgttcagggaagaagaggtgtcgttttgagattctctgggcttcacccctggatg 335

Db 121 GCTGTTCAGGGCAAGAAGGGTGTCTGTTGGAGTTCCTGGGGCTTCACCCCTGGATG 180

Qy 336 ttccaagacacacacctgccagggtttgtggagcaggtgaggtctctgaaggccaaggagt 395

Db 181 TTCCAAGACACACACCTGCCAGGGTGTGTGAGCAGGCTGAGGCTCTGAAGGCCAAGGGAGT 240

Qy 396 ccaggtgtgtgctctcgtcaggttgaatgagtccttctgactggcaggtgggcccagc 455

Db 241 CCAGGTGGTGGCTCTCTGAGTGTAAATGATGCTTGTGACTGGCGAGTGGGGCCGAGC 300

Qy 456 ccacaaggcgaagcaaggttcggctcctggtcgtatcccaactgggcttgggaagga 515

Db 301 CCACAAGCGGAAGCAAGGTTCGGCTCTGCTGCTATCCCACTGGGGCTTTGGGAAGA 360

Qy 516 gacagacttattactagatgattcgtgtgtccatctttgggaatcagctctcaagag 575

Db 361 GACAGACTTATTACTAGATGATTCGCTGGTGTCATCTTTGGGAATCGACGTCTCAAGAG 420

Qy 576 gttctccatggtgtgtacaggatggtcagatggaagccctcgtgaatgtggaaccagatg 635

Db 421 GTTCTCCATGTTGTGTACAGGATGGCATAGTGAAGGCCCTGTAATGTGAACCATGATG 480

Qy 636 aggcctcactcagctcagctggcaccacaatatcatctcagctcgtgagccctgggccaga 695

Db 481 AGGCTCTACCTCGACGCTGGCACCAATATCATCATCAGCTCTGAGGCCCTTGGGCCAGA 540

Qy 696 ttacttctccaccctccctctatctcacctggcagccctgtgctggggccctgcaattg 755

Db 541 TTACTTCTCCACCCCTCCCTATCTCACCTGCCAGCCCTGTGCTGGGGCCCTGCAATTG 600

Qy 756 gaat 759

Db 601 GAAT 604

RESULT 8

BE2976165 1064 bp mRNA EST 22-JAN-2001

LOCUS 602245026F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4336063 5',

DEFINITION mRNA sequence.

ACCESSION BE2976165

VERSION BE2976165.1 GI:12343380

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 1064)

TITLE NIH-MGC http://mgc.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

CDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM1208 row: c column: 08
 High quality sequence stop: 751.
 Location/Qualifiers
 1. .1064
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4336063"
 /clone_lib="NIH_MGC_48"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 274 a 221 c 364 g 205 t

BASE COUNT
 ORIGIN

Query Match 72.7%; Score 585; DB 172; Length 1064;
 Best Local Similarity 100.0%; Pred. No. 1.4e-291;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 ggagcggagtggaagtggcgtggcggtatggagactagctggcgtgtgcccctga 64
 Db 2 GGAGGCGAGTGAAGTGGCGGTGGGGGTATGGGACTAGCTGGCGTGTGGCCCTGA 61
 QY 65 gacgtcagcgggtatatactctcgtggcggtggcggtgcagtctgcggcagcgag 124
 Db 62 GACGCTCAGCGGGGTATATACCTCGGTGGGGCGGGCGGTGAGTCTGCGGCGGCGAG 121
 QY 125 caagacgtgcagtggaagagtggtcgtctgcgggtccgcagttcagcagacgg 184
 Db 122 CAAGCGGTGACGTGAGGAGAGTGGCGTCTGGCGGGTCCGCGAGTTTACGAGCGCG 181
 QY 185 ctgcagccatggcccaatcaagtggagatgcatccacagcagtgagtggtttgaag 244
 Db 182 CTGAGCCATGGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGTGTGTAAG 241
 QY 245 gggagccaggaacaagtgaacctggcagagctgttcaaggcggaagaggtgtgtgt 304
 Db 242 GGGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGGTTGCTGT 301
 QY 305 ttgagttcctggggtcctaccctggatgttccagacacacacctgccaggtttgtg 364
 Db 302 TTGAGTTCCTGGGCGCTTACCCCTCGATGTTCCAGACACACACTGCCAGGGTTGTGG 361
 QY 365 agcaggtcaggtcgtctgaagccaaaggagttccaggtgtgtggtcgtctgagtgtaatg 424
 Db 362 AGCAGGCTGAGGCTCTCAAGGCCAAGGAGTCCAGGTGGTGGCTGTCTGAGTGTAAATG 421
 QY 425 atgctttgtgactggcagtggtggcgagcccaagagcggaagaggttcaggtctcc 484
 Db 422 ATGCCCTTGTGACTGGCGAGTGGGGCGAGCCCAAGAGGGGAAGGTTCCGGCTCC 481
 QY 485 tggctgacccactggggcctttgggaagagagacagacttattactagatgattcgtg 544
 Db 482 TGGCTGATCCACTGGGCGCTTTGGGAGGAGAGACAGACTTATTAAGATGATTCGCTGG 541
 QY 545 tgtccattttgggaatcgacgtctcaagaggttctccatgtgtg 589
 Db 542 TGTCCATCTTTGGGAATCGACGTCTCAAGAGGTTCTCCATGGTGG 586

RESULT 9
 BE410064

LOCUS BE410064 959 bp mRNA EST 21-JUL-2000
 DEFINITION 601300587F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3630548 5',
 mRNA sequence.

ACCESSION BE410064
 VERSION BE410064.1 GI:9346514
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 959)

TITLE NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM318 row: n column: 21
 High quality sequence start: 35
 High quality sequence stop: 792.

Location/Qualifiers
 1. .959
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3630548"
 /clone_lib="NIH_MGC_21"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"

/note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 203 a 255 c 314 g 187 t
 ORIGIN

Query Match 72.5%; Score 584; DB 167; Length 959;
 Best Local Similarity 100.0%; Pred. No. 4.4e-291;
 Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 ggactagctggcgtgtgcccctgagacgtcagcggcgtatatactcgtggggcc 99
 Db 128 GCACCTAGCTGGCGTGTGGCCCTCAGACGCTCAGGGCTATATCTCGTGGGGGCC 187
 QY 100 ggcgttcagctcgcggcagcgagcagcggtgcagtggaagagtggtggcgtctg 159
 Db 188 GCGGTCAGTCTCGCGCAGCGGCGAGAGAGCGGTGCAAGGAGAGAGTGGCGCTGCGC 247
 QY 160 ggggtccgcagtttcagcagagccgctgcagccatgcccccaatcaaggtggagatgcc 219
 Db 248 GGGGTCCGCGAGTTTTCAGCAGAGCCGCTGCAGCCATGCCCCCAATCAAGGTGGGAGATGCC 307
 QY 220 atcccagcagtgagggtgtttgaaggagagcagcggaaggaacaggtggaacgtggagcgtg 279
 Db 308 ATCCACAGCTGAGAGTGTGTAAGGGGAGCCAGGAGACAGGTGAACCTGGCGAGAGCTG 367
 QY 280 ttaaggcgcaagaagggtgtgtgtttgagttcctggggccttacccttcaggtgtgtcc 339
 Db 368 TTCAAGGGCAAGAGGGGTGTCTGTTGGAGTTCCTGGGGCTTTCACCCCTGGATGTTCC 427
 QY 340 aagacacacctgcccaggtttgtgagcaggtcaggtcctgaaggccaaaggagatccag 399


```
Db 534 CTTGGTCCACTCTTTGGGATCGACGCTCAAGAGGTTCTCCATGGTGACAGGATG 593
QY 599 goatagtgagccctgaatgtgaacacagatggcacagggcctcacctgcagctggcac 658
Db 594 GCATAGTGAAGGCCCTGAATGTGAACACAGATGCCACAGGCTCACCTGCAGCCTGGCAC 653
QY 659 ccaatcatcatcacagctcagggccctggccagattactctccacccctccctat 718
Db 654 CCAATATCATCTCACAGCTCTGAGGCCCTGGCCAGATTAATCTCCTCCACCCCTCCCTAT 713
QY 719 ctcaactccagccctcgtcgtgggcccctgcaattggaattgtggcagattctcga 778
Db 714 CTCACCTGCCAGCCCTGCTGCTGGGCCCTGCAATTTGAATGTTGGCAGATTCTTCGAA 773
QY 779 taaacacttggttggtgg 798
Db 774 TAAACACTTGTGTTTGGG 793

RESULT 13
BG325661 710 bp mRNA EST 27-FEB-2001
LOCUS 602424378F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562443 5',
DEFINITION mRNA sequence.
ACCESSION BG325661
VERSION BG325661.1 GI:13132098
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 710)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI274 row: k column: 20
High quality sequence stop: 678.
FEATURES
Location/Qualifiers
1..710
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 174 a 167 c 222 g 144 t 3 others
ORIGIN

Query Match 68.3%; Score 550; DB 152; Length 710;
Best Local Similarity 100.0%; Pred. No. 1.7e-273;
Matches 550; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 ggcgggggtccgcagtttcagcagccgctgcagcattgccccaatcaagggtggagat 216
|||||
```

```
Db 2 GCGGGGGTCCGCGAGTTTTCAGCAGAGCCGCTGCAGCCATGGCCCCAATCAAGGTGGAGAT 61
QY 217 gccatccacagcagtgagggtgtttgaaggaggagccagggaacaagtggaacctggcagag 276
|||||
Db 62 GCCATCCACGACAGTGGAGGTGTTTGAAGGGGAGCCAGGAACAAGGTGAACCTGGCAGAG 121
QY 277 ctgttcaagggaagaagggtgtcgtgttttgaggttcctcctggggcctccaccctcgatgt 336
|||||
Db 122 CTGTTCAAGGGCAAGAGGAGGTGCTGCTGTTGGAGTTTCTGGGGCCCTTCCACCTGGATGT 181
QY 337 tccaaacacacacacgctgcaggggtttgtgagcagcgtgaggtcctcctgaagcccaaggagtc 396
|||||
Db 182 TCCAAACACACACCTGCCAGGGTTTGTGGAGAGGCTGAGGCTCTGAAGGCCAAGGAGTC 241
QY 397 caggtggtggcctgtcgtgagtgtaataatgatgccttctgactggcagtgaggccgagacc 456
|||||
Db 242 CAGGTGTTGGCCTGTCTGAGTGTAAATGATGCCTTTGTGACTGGCGAGTGGGCCGAGCC 301
QY 457 cacaaggcgaaggcaagggttcggctcctcctggctgatacccaactggggcctttgggaaggag 516
|||||
Db 302 CACAAGCGGAGGCAAGGTTCGGCTCCTGGCTGATCCCACTGGGGCCCTTTGGGAAGGAG 361
QY 517 acagacttattactagatgatcgtcgtgttcctcctcctcctcctcctcctcctcctcctcct 576
|||||
Db 362 ACAGACTTATTACTAGATGATTCGCTGGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGG 421
QY 577 ttctccatggtgtacagcagtgatgcatagtgaggccctgaatgtggaaccacagatggcaca 636
|||||
Db 422 TTCTCCATGTTGTACAGGATGCATAGTGAAGGCCCTGAATGTGAACCAACATGGCACA 481
QY 637 ggctcacctgcagcctgcacccaataatcctcctcctcctcctcctcctcctcctcctcctcct 696
|||||
Db 482 GGCTCACCCTGCAGCCTGGCACCCAATATCATCTCACAGCTCTGAGGCCCTGGGCCAGAT 541
QY 697 tacttctctc 706
Db 542 TACTTCTCTCC 551

RESULT 14
BG388478 726 bp mRNA EST 21-JUL-2000
LOCUS 601287129F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3621583 5',
DEFINITION mRNA sequence.
ACCESSION BG388478
VERSION BG388478.1 GI:93333843
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 726)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM295 row: i column: 08
High quality sequence stop: 638.
FEATURES
Location/Qualifiers
1..726
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_44"
/tissue_type="endometrium, adenocarcinoma cell line"
source
```

```

/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
161 a 171 c 233 g 161 t

```

Query Match.	68.1%;	Score 548;	DB 167;	Length 726;
Best Local Similarity	100.0%;	Prod. No. 1.9e-272;		
Matches 548;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	144	agagtggagctctgagcggtccacagtttcacagagcgcctcagccatggccccaat	203	
Db	1	AGAGTGGGCGTCTGGCGGGTCCGAGTTTCAGCAGAGCCGCTCAGGCATGGCCCCAAT	60	
QY	204	caaggtggagatgccatccacagcagtggaagtgtttgaaggagagccagggaaacaagt	263	
Db	61	CAAGGTGGAGATGCCATCCAGCAGTGGAGGTCTTTGAAGGGAGCCAGGGAACAAGGT	120	
QY	264	gaacctggcagagctgttccaagggaagaggggtgtcgtctttggagtctcctggggcctt	323	
Db	121	GAACCTGGCAGAGCTGTCAAGGGCAAGAGGGTGTGCTGTTTGGAGTTCCTGGGGCCTT	180	
QY	324	cacccttgatgttccaaagacacacacctgcacaggttttgtgagcaggtgaggtctctgaa	383	
Db	181	CACCCCTGGATGTTCCAAGACACACCTGCCAGGGTTGTGGAGCAGGCTGAGGCTCTGAA	240	
QY	384	ggccaaaggagtcaggtggtggtcgtctgagtggttaatgatcctttgtgactggcga	443	
Db	241	GGCCAAAGGGAGTCAGGTGGTGGCTGTCTCAGTGTTTAATGATGCTTGTGACTGGCGA	300	
QY	444	gtggggtccagaccacaaagcggaagcaaggttcggtcctggtcgtatccactggggc	503	
Db	301	GTGGGGCGGACCCACAAAGGGGAAGGCAAGGTTCCGGCTCTGGCTGATCCCATGGGGC	360	
QY	504	ctttgggaagagacagacttattactagatgattcgtggtgccatctttgggaatcg	563	
Db	361	CTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGTGTGTGCCATCTTTGGGAATCG	420	
QY	564	acgtctcaagaggttctccatggttggttacagatggcatagtgaagccctgaatggga	623	
Db	421	ACGTCTCAAGAGGTTCTCCATGTTGGTACAGGATGGCATAGTGAAGGCCCTGGAATGTGA	480	
QY	624	accagatggcacagggcctcacctcagcctggcaccacaatatcatctcacagctctgag	683	
Db	481	ACCAGATGGCACAGGCTCACCTCAGCCTGGCACCCCAATATCATCTCACAGCTCTGAGG	540	
QY	684	ccctgggc	691	
Db	541	CCCTGGGC	548	

RESULT 15	BE394469	582 bp	mRNA	EST	21-JUL-2000
LOCUS	BE394469				
DEFINITION	50131117581 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:363290 5', mRNA sequence.				
ACCESSION	BE394469				
VERSION	BE394469				
KEYWORDS	BE394469.1	GI:9339834			
SOURCE	EST.				
ORGANISM	human.				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 582)				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				

Mon Oct 22 14:46:50 2001

us-09-486-167a-1.oligo.rst

Page 13

Search completed: October 22, 2001, 13:21:17
Job time: 4142 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2001, 10:41:25 ; Search time 1294.87 seconds
(without alignments)
5876.688 Million cell updates/sec

Title: US-09-486-l67A-1
Perfect score: 805
Sequence: 1 gccaggagcgagtggaag.....tltggtttgcggaaaaaaa 805

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: gb_est44:*
45: gb_est45:*
46: gb_est46:*
47: gb_est47:*

em_esthum10:*
44: em_esthum11:*
45: em_esthum12:*
46: em_esthum13:*
47: em_esthum14:*
48: em_esthum15:*
49: em_esthum16:*
50: em_esthum17:*
51: em_esthum18:*
52: em_esthum19:*
53: em_esthum20:*
54: em_esthum21:*
55: em_esthum22:*
56: em_esthum23:*
57: em_esthum24:*
58: em_esthum25:*
59: em_esthum26:*
60: em_esthum27:*
61: em_esthum28:*
62: em_estin1:*
63: em_estin2:*
64: em_estin3:*
65: em_estin4:*
66: em_estin5:*
67: em_estom1:*
68: em_estom2:*
69: em_estov1:*
70: em_estov2:*
71: em_estpl1:*
72: em_estpl2:*
73: em_estpl3:*
74: em_estpl4:*
75: em_estpl5:*
76: em_estpl6:*
77: em_estpl7:*
78: em_estpl8:*
79: em_estpl9:*
80: em_estpl10:*
81: em_estro1:*
82: em_estro2:*
83: em_estro3:*
84: em_estro4:*
85: em_estro5:*
86: em_estro6:*
87: em_estro7:*
88: em_estro8:*
89: em_estro9:*
90: em_estro10:*
91: em_estro11:*
92: em_estro12:*
93: em_estro13:*
94: em_estro14:*
95: em_estro15:*
96: em_estro16:*
97: em_estro17:*
98: em_estro18:*
99: em_estro19:*
100: em_estro20:*
101: gb_est25:*
102: gb_est26:*
103: gb_est27:*
104: gb_est28:*
105: gb_est29:*
106: gb_est30:*
107: gb_est31:*
108: gb_est32:*
109: gb_est33:*
110: gb_est34:*
111: gb_est35:*
112: gb_est36:*
113: gb_est37:*
114: gb_est38:*
115: gb_est39:*
116: gb_est40:*
117: gb_est41:*
118: gb_est42:*
119: gb_est43:*
120: gb_est44:*
121: gb_est45:*
122: gb_est46:*
123: gb_est47:*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

```

FEATURES
source
Location/Qualifiers
1. .918
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE007YL04"
/clone_lib="LTI_FL002_PL1"
/lab_host="DH10B"
/note="Organ: Placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
218 a 225 c 296 g 178 t 1 others
BASE COUNT
ORIGIN

```

Query Match	99.2%;	Score 798.6;	DB 106;	Length 918;
Best Local Similarity	99.5%;	-Pred. No. 8.7e-197;		
Matches 801: Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

60	Qy	1	gccaggaagcggagtggaagtggccgtggggcgggtatggagctactgactgcgctgcgcgc	60
62	Db	62	gccagggagcggagtggaagtggccgtggggcgggtatggagctactgactgcgctgcgcgc	120
61	Qy	61	ctgagacgctcaagcgggctatatactactgctcggtggggcggcgggtcagctctgcggcagcg	120
122	Db	122	ctgagacgctcagcgggctatatactactgctcggtggggcggcgggtcagctctgcggcagcg	180
121	Qy	121	gcagcaagacggtgcagtgaaagagagtgaggtctctggcggggtccgcagtttcaagcaga	240
182	Db	182	gcagcaagacggtgcagtgaaagagagtgaggtctctggcggggtccgcagtttcaagcaga	240
181	Qy	181	gcgctgcagccatggccccaatcaaggtgggagatgccatcccagcagtggaagtgctt	240
242	Db	242	gcgctgcagccatggccccaatcaaggtgggagatgccatcccagcagtggaagtgctt	300
241	Qy	241	gaaggggagccaggaaacaaagtgaaacctggcagagctgttcaaaggcgaagaaggggtgtg	300
302	Db	302	gaaggggagccaggaaacaaagtgaaacctggcagagctgttcaaaggcgaagaaggggtgtg	360
301	Qy	301	ctgtttggagttctcggggccctcacccctggatgttccaagaacacacctgcagggttt	360
362	Db	362	ctgtttggagttctcggggccctcacccctggatgttccaagaacacacctgcagggttt	420

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	798.6	99.2	918	106	AL541916	AL541916	AL541916
2	773.4	96.1	818	106	AL535869	AL535869	AL535869
3	766.8	95.3	850	105	AL520418	AL520418	AL520418
4	715.2	88.8	831	154	BG498369	BG498369	BG498369
5	711	88.3	825	155	BG545942	BG545942	BG545942
6	705.2	87.6	889	172	BG482283	AL520417	AL520417
7	701.8	87.1	741	105	BF975906	BF975906	BF975906
8	692.6	86.0	790	140	BE784530	BE784530	BE784530
9	689.8	85.7	1064	172	BF976165	BF976165	BF976165
10	674.4	83.8	703	175	BG287024	BG287024	BG287024
11	673.4	83.7	760	154	BG482283	BG482283	BG482283
12	669	83.1	959	167	BG410064	BE410064	BE410064
13	668.6	83.1	840	172	BG026788	BG026788	BG026788
14	667	82.9	677	32	AV650400	AV650400	AV650400
15	666.4	82.8	842	140	BE796516	BE796516	BE796516
16	663.6	82.4	813	153	BG438079	BG438079	BG438079
17	660.6	82.1	734	106	AL541915	AL541915	AL541915
18	652.2	81.0	751	165	BE300398	BE300398	BE300398
19	650.6	80.8	779	139	BE728013	BE728013	BE728013
20	649.8	80.7	909	139	BE735282	BE735282	BE735282
21	647.2	80.4	954	174	BG167833	BG167833	BG167833
22	645.6	80.2	794	139	BE733545	BE733545	BE733545
23	644.4	80.0	713	154	BG470793	BG470793	BG470793
24	642.8	79.9	836	152	BG325661	BG325661	BG325661
25	638.2	79.3	836	140	BE795522	BE795522	BE795522
26	633.2	78.7	851	172	BF975982	BF975982	BF975982
27	632	78.5	1029	139	BE736544	BE736544	BE736544
28	630.2	78.3	757	140	BE787187	BE787187	BE787187
29	630.2	78.3	845	137	BE562997	BE562997	BE562997
30	625	77.6	794	174	BG113765	BG113765	BG113765
31	624.4	77.6	768	139	BE745244	BE745244	BE745244
32	622.2	77.3	787	173	BG109369	BG109369	BG109369
33	622	77.3	622	175	BG325577	BG325577	BG325577
34	620.4	77.1	918	150	BE530246	BE530246	BE530246
35	620.2	77.0	910	153	BG403357	BG403357	BG403357
36	616.2	76.5	746	174	BG166688	BG166688	BG166688
37	616.2	76.5	758	168	BF679207	BF679207	BF679207
38	615.4	76.4	874	175	BG251461	BG251461	BG251461
39	613.4	76.2	811	141	BE907261	BE907261	BE907261
40	611.4	76.0	1040	173	BG105407	BG105407	BG105407
41	610.2	75.8	911	174	BG119624	BG119624	BG119624
42	609.6	75.7	898	140	BE780877	BE780877	BE780877
43	609.4	75.7	665	141	BE893687	BE893687	BE893687
44	607.8	75.5	942	175	BG333661	BG333661	BG333661
45	606.8	75.4	858	152	BE6237084	BE6237084	BE6237084
46	604.3	75.4	942	175	BG3337084	BG3337084	BG3337084

ALIGNMENTS

RESULT	1
AL541916	
LOCUS	AL541916 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE007YL04 5 prime
DEFINITION	AL541916 918 bp mRNA EST 16-FEB-2001
	, mRNA sequence.
ACCESSION	AL541916
VERSION	AL541916.1 GI:12873445
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 918)
AUTHORS	Li,W.B., Gruber,C., Jesse,J., Polayes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope

Db	180	TGCAGCCAWGGCCCCCAATCAAGGTGGGAGATGCCATCCACGACGTGGAGGTGTTTGAAGG	239
QY	246	ggagccagggaacaaggtgaacctggcagagctgttcaagggaacaagggtgctgtgtt	305
Db	240	GGAGCCAGGGAAACAAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAAGGTGTGCTGT	299
QY	306	ttgagttctctggggccttcacccctggatgttccaaagcacacactgcacagggtttgtga	365
Db	300	TGGAGTTCCTGGGGCTTCACCCCTGGATGTTCCAGACACACCTGCCAGGTTTGTGGA	359
QY	366	gcagctgagagctctgaaggccaaggagtcacaggttggtgacctgtctgagttaatga	425
Db	360	CGAGGCTGAGGCTCTGAAGGCGCAAGGAGGTCCAGGTGGTGGCTGTCTGAGTGTAAATGA	419
QY	426	tgctttgtgactgagcagtggggcgcagcccaagcgcggaagcgaaggttcggctcct	485
Db	420	TGCCTTTGTGACTGGCAGTGGGGCCGAGCCCAAGGCGAAGGTCAGGTTCGGCTCCT	479
QY	486	ggctgataccactggggccctttgggaaggagacagacttattactagatgattcgtcgtg	545
Db	480	GGCTGATCCCACTGGGGCCCTTTGGGAAGGAGACAGACTTATTACTAGATGATTTCGTGT	539
QY	546	gtcca-cttttgggaatcgacgtctcaaaaggttccatggtgtggtacaggatggcatag	604
Db	540	GTCCACTCTTTGGGAATCGACGCTCTCAAGAGGTTCTCCATGGTGTACAGGATGGCATAG	599
QY	606	tgaaggccctgaaatgtggaaccagatggcacagggcctcacctgcagcctggcacccaata	664
Db	600	TGAAGGCCCTGAATGTGGAACCAAGATGGCAGCGCCTCACCTGCGCTGGCACCAATA	659
QY	666	tcaatcacagctctgagggccctggggccagattacttcacacccctccctatctcacc	724
Db	660	TCATCTCACAGCTCTGAGGCCCCCTGGGCCAGATTACTTCCACGCCCTCCCTATCTCACC	719
QY	726	tggccagccctgtctggggccctgcaattggaattgtggcagattttctgcaataaaca	784
Db	720	TGCCACGCCCTGTGCTGGGGCCCTGCAATTGGAATGTGGCCAGATTTCTGCAATAACA	779
QY	786	cttggtgttcgggaaaaaaa 805	
Db	780	CTTGTTGTTTGGCGCCAAAAA 800	
RESULT	3		
AL520418			
LOCUS	AL520418	850 bp	mRNA
DEFINITION	AL520418	LRI_NFL004_NBC2	Homo sapiens cDNA clone CS0DB006YE21
ACCESSION	AL520418		
VERSION	AL520418		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Li, W.B., Gruber, C., Jessee, J., and Pollayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
CONTACT	Contact: Genoscope		
GENOSCOPE	Genoscope - Centre National de Sequencage		
BP	BP 191 91006 EVRY cedex - France		
EMAIL	Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
FEATURES	Location/Qualifiers		
SOURCE	1. 850		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DB006YE21"		
	/clone_lib="LRI_NFL004_NBC2"		
	/sex="male"		
	/tissue.type="neuroblastoma cells"		
	/lab_host="DH10B"		
	/note="organism: brain; vector: scudcpom 6.0; insert: 850bp"		

was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT	185 a	212 c	280 g	170 t	3 others
ORIGIN					
Query Match	95.3%; Score 766.8; DB 105; Length 850;				
Best Local Similarity	98.4%; Pred. No. 1.5e-188;				
Matches	793; Conservative	2; Mismatches	9; Indels	2; Gaps	2;
QY	1	gccagagcgagctgaagtggccgtggcggtgtagtgaactagctgctgctgcgc	60		
Db	27	gccagagcgagctgaagtggccgtggcggtgtagtgaactagctgctgctgcgc	86		
QY	61	ctgagacgctcagcggtctatactctgctggtggcggtgtagtgaactagctgctg	120		
Db	87	ctgaaacgctcaacggctatatactctgctggtggcggtgtagtgaactagctgctg	146		
QY	121	gcagcaagacgctgtagtgaagagagtggtgctgctggtggcggtgtagtgaactagctgctg	180		
Db	147	gcagcaagacgctgtagtgaagagagtggtgctgctggtggcggtgtagtgaactagctgctg	206		
QY	181	gccctgcagcctgagcccaatacaagtggagagtggtgctgctggtggcggtgtagtgaactagctgctg	240		
Db	207	gccctgcagcctgagcccaatacaagtggagagtggtgctgctggtggcggtgtagtgaactagctgctg	266		
QY	241	gaagggagcgaggaagagtggaactggtgagagtggtgctgctggtggcggtgtagtgaactagctgctg	300		
Db	267	gaagggagcgaggaagagtggaactggtgagagtggtgctgctggtggcggtgtagtgaactagctgctg	326		
QY	301	ctgtttgagttcctggggtcctcaccctggatgtgttcccaagacacacacctgccagggttt	360		
Db	327	ctgtttgagttcctggggtcctcaccctggatgtgttcccaagacacacacctgccagggttt	386		
QY	361	gtgagacgctgaggtctgaaagcccaagggagtcaggtggtggcggtgtagtgaactagctgctg	420		
Db	387	gtgagacgctgaggtctgaaagcccaagggagtcaggtggtggcggtgtagtgaactagctgctg	446		
QY	421	aatgatgcttctgactggcgagtggtggcgagccacacagcggaagcaaggttcgg	480		
Db	447	aatgatgcttctgactggcgagtggtggcgagccacacagcggaagcaaggttcgg	506		
QY	481	ctcctggctgacccactggggtccttgggaagagagacagacttattactagatgctg	540		
Db	507	ctcctggctgacccactggggtccttgggaagagagacagacttattactagatgctg	566		
QY	541	ctgtgttcactcttgggaatcgacgtctcaagaggttctccatggttggtacaggaatggc	600		
Db	567	ctgtgttcactcttgggaatcgacgtctcaagaggttctccatggttggtacaggaatggc	625		
QY	601	atagtgaagccctgaatgtggaaccagatggcacagcctcactcagcctggcaccc	660		
Db	626	atagtgaagccctgaatgtggaaccagatggcacagcctcactcagcctggcaccc	685		
QY	661	aatatcactcacagctctgagggcctgggacagatcttctccacccctccctatct	720		
Db	686	aatatcactcacagctctgagggcctgggacagatcttctccacccctccctatct	745		
QY	721	cacctgcccagcctctgctggggccctgcaatt-gaatgtgtggccagatattctgcaat	779		
Db	746	cacctgcccagcctctgctggggccctgcaatt-gaatgtgtggccagatattctgcaat	805		
QY	780	aaacactgtgtgttcggaataaaaa	805		
Db	806	aaacactgtgtgttcggaataaaaa	831		

RESULT	4
LOCUS	BG498369
DEFINITION	602544160F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:466430 5', mRNA sequence.
ACCESSION	BG498369
VERSION	BG498369.1
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 831)
TITLE	NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-r@mail.nih.gov Tissue Procurement: DCTD/DP cDNA Library Preparation: CLONETECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM1473 row: h column: 15 High quality sequence stop: 775.
FEATURES	Location/Qualifiers
Source	1..831
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:466430"
	/clone.lib="NIH_MGC_60"
	/tissue.type="adenocarcinoma"
	/lab_host="DH10B (T1 phage-resistant)"
	/note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site_1: SfII (ggcgctcgcc); Site_2: SfII (ggccatctggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCGGCGGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."
BASE COUNT	180 a 195 c 282 g 173 t 1 others
ORIGIN	
Query Match	88.8%; Score 715.2; DB 154; Length 831;
Best Local Similarity	97.7%; Pred. No. 3.8e-175;
Matches	757; Conservative 0; Mismatches 14; Indels 4; Gaps 3;
QY	26
Db	2
QY	86
Db	62
QY	146
Db	122
QY	206
Db	182
QY	266

Db 601 CCTGAATGTGAACAGATGCCACAGCCTCA-CTGACGCTGCACCCATATCACTC 659
Qy 672 acagctctgagccctggcagattactctctccacccctccctctctctctccag 731
Db 660 ACAGCTCTGAGGGCTGGGCCAGATTAATCTCTCAACCTT-CTATCTCA-CTGGCCAG 717
Qy 732 cccctgtctggggcc 746
Db 718 GCCTGTCTGGGGCC 732

RESULT 9
BF976165 1064 bp mRNA EST 22-JAN-2001
LOCUS 602245026F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4336063 5',
DEFINITION mRNA sequence.
ACCESSION BF976165
VERSION BF976165.1 GI:12343380
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1064)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM1208 row: c column: 08
High quality sequence stop: 751.
Location/Qualifiers
1..1064
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4336063"
/clone_lib="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;
Site: 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 274 a 221 c 364 g 205 t
ORIGIN

Query Match 85.7%; Score 689.8; DB 172; Length 1064;
Best Local Similarity 95.4%; Pred. No. 1.6e-168;
Matches 764; Conservative 0; Mismatches 32; Indels 5; Gaps 5;

Qy 5 gaggcgcagtgagtgccgtggcgggtatggactagctgcgtgtgcgcctga 64
Db 2 GGAGCGGAGTGAAGTGGCGCGGTATGGACTAGCTGGCGTGGCCCTGA 61
Qy 65 gacgctcagcggcgtatctactcgtgtggcggcggcgtcagtcggcagcgccag 124
Db 62 GACGCTCAGCGGGCTATATCTCTGCTGGTGGCGCGCGTCTGCTGGCGCGCAG 121
Qy 125 caagacggtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 184

Db 122 CAAGACGGTGCAGTGAAGGAGAGTGGGCGCTCTGGCGGGTCCGAGTTTCACGACGCCG 181
Qy 185 ctgacgcatgcccacaaatcaagtgagatgcccacatccacagtgagggtgtttgaag 244
Db 182 CTCAGCCATGGGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGAAG 241
Qy 245 gggagcaggggaacaaagggtgaacctggcagagctgttcaagggaagaggggtgtgtctg 304
Db 242 GGGAGCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGGGTGTCTGT 301
Qy 305 ttggagttcctggggccttcacccctggatgtttccaaagacacacacacaggtttgtg 364
Db 302 TTGAGTTCCTGGGGCTTCAACCCCTGGATGTTTCCAAAGACACACCTGGCAGGGTGTGG 361
Qy 365 agcaggtcgtgagctctgaagccaagggagtcaggtggcctgtctgagtggttaag 424
Db 362 AGCAGGCTGAGGCTCTGAAGGCCAAGGGAGTCCAGGTGGTGGCTGTCTGAGTGTAAATG 421
Qy 425 atgcctttgtgactggcagtggtggggccgagccacacagcgggaagcgaaggttcggctcc 484
Db 422 ATGCTTTGTGACTGGCGAGTGGGGCGCGAGCCACACAGCGGGAAGCAAGGTTTCGGCTCC 481
Qy 485 tggctgacccactggggcctttggaagagagagacagacttatctactagatgattcgtcg 544
Db 482 TGGCTGATCCACATGGGGCCCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCTGG 541
Qy 545 tgcctctttgggaatcgacgtctcgaagaggttctccatggtt-ggtacagggatggcata 603
Db 542 TGTCCATCTTTGGGAATCGACGCTCTCAAGAGGTTCTCCATGGTGGGTACAGGATGGCATA 601
Qy 604 gtgaagccctgaatgtggaaacagatggcacagggcctcaccctgcagcctggcaccat 663
Db 602 GTGAAGG-CCTGAATGTGGAACACAGATGGCACAGGCGCTCAGCTGCAGCTGGCACCAAT 660
Qy 664 atcatctcagcctcagcctggcggcagattactctccaccctccctctatctcac 723
Db 661 ATCATCTCAGACTCTCTGAGG-CCITGGGCCAGATTACTTCTTACA-CCCTCCCTATCTCAG 718
Qy 724 ctgcccagccctgtgctggggccctcgaattggaattgtggccagatttctgcaataaac 783
Db 719 CTGCCCAGCCCTGTGCTGGGGCGCTGCATTGGAATG-TGGGCAGATTCTCGGATAAAC 777
Qy 784 acttggtgttgccggaataaaa 804
Db 778 TTGTGTTGGGGCAAAAAA 798

RESULT 10
BF976165
LOCUS 602382319F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4499716 5',
DEFINITION mRNA sequence.
ACCESSION BF976165
VERSION BF976165.1 GI:13040451
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 703)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLAM10364 row: f column: 05
High quality sequence stop: 702.

FEATURES
source Location/Qualifiers
1. .703

FEATURES	LOCUS
source	DEFINITION
1. .703	ACCESSION
/organism="Homo sapiens"	VERSION
/db_xref="taxon:9606"	KEYWORDS
/clone="IMAGE:4499716"	SOURCE
/clone_lib="NIH_MGC_93"	ORGANISM
/tissue_type="transitional cell papilloma, cell line"	REFERENCE
/note="Organ: bladder; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."	AUTHORS
142 a 177 c 234 g 150 t	TITLE
	JOURNAL
	COMMENT

Query Match 83.8%; Score 674.4; DB 175; Length 703;
Best Local Similarity 99.6%; Pred. No. 1.4e-164;
Matches 697; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY	89	tcggtgggcccgggtcagttctgcggcagcggcagcaagacgtgacgtgaaggagagt	148
Db	1	TCGGTGGGCGCGCGGTCAgTCTCGCGCAGCGCAGCAAGACGGTGCAGTGAAGGAGACT	60
QY	149	ggcgctctggcggggTccgcagtttcacagacagcagctgcagccatggcccacaatacagg	208
Db	61	GGGCGTCTGGCGGGTCCGcAGTTTCAGCAGAGCCGCTGCAGCCATGGCCCCCAATCAAGG	120
QY	209	tgggaatgccatcccagcagtgagaggtgtttgaaggggagccaggaacaaggTgaacc	268
Db	121	TGGAGATGCCATCCAGCAGTGGAGTGTtTGAAGGGAGCCAGGGAACAGGTGAACC	180
QY	269	tggcagacgtgttcaaggccaagaagggtgctgttttgaggttccctggggccctcaccc	328
Db	181	TGGCAGAGCTGTtCAAGGCAAGAGGGTGTCTGTtTGGAGTCTCTGGGGCTTTCACC	240
QY	329	ctggatgtttccaagacacacctgccagggtttgttgagcagagctgaggtctgaaggcca	388
Db	241	CTGGATGTtTCCAAGACACACCTGCCAGGTTTGTGSAGCAGGCTCAGGCTCTGAAGGCCA	300
QY	389	agggaatccagatgtggcctctctgagttgtaataatgactcttctgactggcagtgagg	448
Db	301	AGGGAGTCCAGTGTGGCTCTGtGAGTGTtTAATGATGCCCTTGTGTGACTGGCGAGTGGG	360
QY	449	gcccagccacaaggcggaaaggcaaggttcggctcctggctgactccactggggccttgg	508
Db	361	GCCGAGCCACAAAGCGGAAGGCAAGGTTGGCTTCTGGCTGATCCCACTGGGGCTTTG	420
QY	509	ggaaggagacagacttatactagatgttcgctgggtgccactctttggaaatcgacgtc	568
Db	421	GGAAAGAGACAGACTTATtACTAGATGATCGCTGGTGTCTCCTCTTTGGGAATCGACGTC	480
QY	569	tcaagagttctccatgggtgttacaggatggcatagtgaaggccctgaatggaacag	628
Db	481	TCAAGAGGTTCTCCATGCTGTACAGGATGGCATAGTGAAGGCCCTGAATGTGAACCAAG	540
QY	629	atggcacaggctcaactgcagcctggcacccaataatactcacagctctgaggccctg	688
Db	541	ATGGCACAGGCTCACTTGCAGCTGGCACCCAATATCATCTCACAGCTCTGAGGCCCTG	600
QY	689	ggccagattacttctccacccctcccttatctcactgccagccctgtctgagggccct	748
Db	601	GGCCAGATTACTTCTCCA-CCCTCCCTATCTCACTTGCCAGCCGGTGTGCTGGGGCCCT	659
QY	749	gcaa-ttggaagtggccagattctgcataaacactt	787
Db	660	GCAATTTGGAATGTTGGCAGATTCTTGCATAAACACTT	699

RESULT 11

BG482283	760 bp	EST	21-MAR-2001		
LOCUS	603526614F1 NIH_MGC_21	EST	IMAGE:4650363 5',		
DEFINITION	mRNA sequence.				
ACCESSION	BG482283				
VERSION	BG482283.1	GI:13414562			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1. (bases 1 to 760)				
TITLE	NIH-MGC http://mgc.nci.nih.gov/				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapbs-remail.nih.gov				
	Tissue Procurement: ATCC				
	CDNA Library Preparation: Ling Hong/Rubin Laboratory				
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
	DNA Sequencing by: Incyte Genomics, Inc.				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LLNL at:				
	http://image.llnl.gov				
	Plate: LfCM1431	row: k	column: 04		
	High quality sequence stop:759.				

FEATURES	source
----------	--------

```

1: 7900
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4650363"
/clone_lib="NIH_MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: placenta; Vector: pORF7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).".
144 a 193 c 273 q 150 t
BASE COUNT

```

BASE COUNT
ORIGIN

Query Match 83.7%; Score 673.4; DB 154; Length 760;
Best Local Similarity 99.4%; Pred. No. 2.7e-164;
Matches 707; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

QY	1	gcagagagcgagtggaagtcgcgcctggggcggggtatggactagctggcgtgtgcgcc	60
Db	49	GCAGAGAGCGGAGTCGGAAGTCGCCCTGGGCGGGGTATGGACTAGCTGGCGTGTGCGCC	108
QY	61	ctgagacgctcagcgggctatactcgttcggtggggccgcggtcagctctcgcgcagcg	120
Db	109	CTGAGACGCTCAGCGGGCTATATACTCGTGGGGCGGGCTCAGTCTCGGCACGCC	168
QY	121	gcagcaagacggtgcagtgaaagagagtgggcgctctgcggggctcgcgaatttcagcaga	180
Db	169	GCAGCAGAGCGGTGCAGTAAGGAGAGTGGGGCGTCTGGCGGGGTCCGAGTTTCAGCAGA	228
QY	181	gccgctgcagccatggccccaatcaaggtggggagatgccatcccgacgagtggagtggtt	240
Db	229	GCCGCTGACGCCATGCCCCCAATCAAGGTGGGAGATGCCATCCACGAGTGGAGTGTTT	288
QY	241	gaaggagcgacaggaacaaagtgaaacctggcagagctgttcaaggggcaagaagggtgtg	300
Db	289	GAAAGGGGACCCAGGGAACAAGGTGAACCTTGGCAGAGCTGTTCAGGGGCAAGAAGGGTGTG	348
QY	301	ctgtttgagttctctggggccctccacctggatgttccaagacacacctgcacagggttt	360
Db	349	CTGTTTGAAGTCTCTGGGGCCCTTCAACCGTGGATCTTCCAAGACACACTCGCAGGGTTT	408

QY 361 gtgagacagctgagctctgaagccaaaggaggtccaggtggtgctgtctgagttt 420
 Db 409 GTGGAGCAGCTGAGCTCTGAAGGCCAAAGGAGTCCAGGTGGTGGCCCTGTCTGAGTGT 468
 QY 421 aatgatgctcttctgactgagctgagtgccgagcccaagagcgaaggaaggttcg 480
 Db 469 AATGATGCTCTTGTGACTGCGAGTGGGCGGAGCCACAAAGCGGAAGGAGTTCGG 528
 QY 481 ctctggtgatccactggggtcttggg-aaggagacagacttattactagatgattc 539
 Db 529 CTCCTGCTGATCCACTGGGCGCTTTGGGCAAGGAGACAGACTTATTACTAGATGATTC 588
 QY 540 gctggttccatcttgggaatcagctcgaagagttctccatggtgtacagatgg 599
 Db 589 GCTGGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTTCTCCATGGGTACAGATGG 648
 QY 600 catagtgaagccctgaatgtgaaccagatggcagggcctcacctgagcctggcacc 659
 Db 649 CATAGTGAAGCCCTGAATGTGAACCCAGATGGCACAGGCTCACCTGCAGCCTGGCA-C 707
 QY 660 caatatactcagctctgagccctgggcccagagattacttctccacc 710
 Db 708 CAATATCATCTCAG-TCGAGGGCCTGGGCCAGATTACTTCTCCACC 757

RESULT 12
 BE410064 959 bp mRNA EST 21-JUL-2000
 LOCUS 601300587F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3630548 5',
 DEFINITION mRNA sequence.
 ACCESSION BE410064
 VERSION BE410064.1 GI:9346514
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 959)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: L1CM318 row: n column: 21
 High quality sequence start: 35
 High quality sequence stop: 792.
 Location/Qualifiers
 1..959
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3630548"
 /clone_lib="NIH_MGC_21"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; CDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 203 a 255 c 314 g 187 t

FEATURES
source

RESULT 13
 BG026788 840 bp mRNA EST 24-JAN-2001
 LOCUS 602294255F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4388940 5',
 DEFINITION mRNA sequence.
 ACCESSION BG026788
 VERSION BG026788.1 GI:12414756
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 840)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

Query Match 83.1%; Score 669; DB 167; Length 959;
 Best Local Similarity 94.9%; Pred. No. 3.9e-163;
 Matches 713; Conservative 0; Mismatches 35; Indels 3; Gaps 2;
 QY 1 gccagagagcggagtggaagtggccgtggggcggtatgggactagctggcgctgagcc 60
 Db 89 GCCAGAGCGGAGTGGAAAGTGGCGGTGGGGCGGTAATTGGACTAGCTGGCGTGTGCC 148
 QY 61 ctgagacactcagcgggctatatctcgtggtgggcccggcggtcagctctcgcgagcg 120
 Db 149 CTGACAGCTAGCGGGCTATATATCTGCTGGTGGGGCGCGGTTCAGTCTCGCGCAG 208
 QY 121 gcagcaagacagtggtcagtggaaggagagtggtggcgctggtgggggtccgcagttcagcaga 180
 Db 209 GCAGCAAGACGGTGCAGTGAAGGAGTGGCGTCTGGCGGGGTCCGAGGTTTCAGCAGA 268
 QY 181 gccgtgcagccatggcccccaatcaagtggtggagatgccatcccgagtgaggtgtt 240
 Db 269 GCCGTGACGCCATGCCCCCAATCAAGGTGGGAGATGCCATCCACAGCAGTGGAGGTGTT 328
 QY 241 gaaggggagccaggggaacaagtggaacctggcagagctgttcaagggttcaaggaggtgtg 300
 Db 329 GAAGGGAGCCAGGGAACAAGCTGAACCTGGCGAGAGTGTTCAGGGCAAGAAGGTGTG 388
 QY 301 ctgtttggagttcctggggcctccacccctgagtttccaaagacacacacgtccaggttt 360
 Db 389 CTGTTTGGAGTTCTCTGGGGCTTTCACCCCTGGATGTTCGAAGACACACCTGCGAGGGT 448
 QY 361 gtggagcagctgaggtctgaagccaaaggaggtggtggtggtggtggtggtggtggt 420
 Db 449 GTGGAGCAGCTGAGGCTCTGAAGGCCAAAGGAGTCCAGGTGGTGGTGGTGGTGGTGT 508
 QY 421 aatgatgctcttctgactgagtggtgggcccagagcccaagggcgaaggaggttcgg 480
 Db 509 AATGATGCTCTTGTGACTGCGAGTGGGCGCGAGCCACAAAGCGGAAGGAGTTCGG 568
 QY 481 ctctggtgatccactggggtctggaagggagagagactattactagatgattc 540
 Db 569 CTCTGGCTGATCCACTGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCG 628
 QY 541 ctggttccatcttgggaatcgacgtctcaagaggttctccatggtggtacagatggc 600
 Db 629 CTGCTGTCATCTTTGGGAATCGACGCTCAAGAGTTCCTCCATGTTGTACAGGATGC 688
 QY 601 atagtgaagccctgaatgtgggaaccagatggcacaggcctacactgcagcctggcacc 660
 Db 689 ATAGTGAAGGCCCTGAATGTGG-ACCAGATGGCACAGGCTCACCTGCGAGCCTGGCACC 747
 QY 661 aatatactct--cacagctcagggccctgggcccagattacttctccaccctccctat 718
 Db 748 AATATCATCTCAACAGTCTGAGGCGCTTGGGCCAGATTACTTCTTACAGCCCATTCCT 807
 QY 719 ctcacctggccagccctgtgtgtggggccctg 749
 Db 808 ATTTTCACCTTGGCAAGGCTTGTTCATG 838

RESULT 13
 BG026788 840 bp mRNA EST 24-JAN-2001
 LOCUS 602294255F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4388940 5',
 DEFINITION mRNA sequence.
 ACCESSION BG026788
 VERSION BG026788.1 GI:12414756
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 840)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue procurement: ATCC
CDNA library Preparation: Life Technologies, Inc.
CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10075 row: n column: 13
High quality sequence stop: 720.

FEATURES
SOURCE

```

mg3: query sequence scop: 720.
location/Qualifiers
  1. 840
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:438940"
    /clone.lib="NIH_MGC_86"
    /tissue_type="osteosarcoma, cell line"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: bone; vector: PCMW-SPORT6; Site_1: NotI;
    Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
    Average insert size 1.533 kb. Library enriched for
    full-length clones and constructed by Life Technologies.
    Note: this is a NIH_MGC Library."
  189 a 198 c 287 g 166 t

```

	BASE COUNT
	ORIGIN

```

BASE COUNT      189 a      198 c      287 g      166 t
ORIGIN          note: ttns is a NIH_MGC Library..

Query Match      83.1%      Score 668.6;  DB 172;  Length 840;
Best Local Similarity 93.7%;  Prod. No. 4.8e-163;
Matches 754;  Conservative 0;  Mismatches 39;  Indels 12;  Gaps 5;

```

Query Match 83.1%; Score 668.6; DB 172; Length 840;
Best Local Similarity 93.7%; Pred. No. 4.8e-163;
Matches 754; Conservative 0; Mismatches 39; Indels 12; Gaps 5;

QY 1 gccaggaggcggaftggaagtggcagtgggcggtatggagactagctggcgtgtgagcc 60

8 G C C A G G A G C G G A G T G G A A G T G G C C G G T A T G G G A C T A C T G G C G T G T G C G C C 67

Qy 61 C T G A G A C G C T C A G C G G G T A T A C T C G T G G G G C G G C G G T C A G T C A G T C G C G G C A G C G 120

68 CTGAGACGCTCAGGGGCTATATACTCGTGGTGGGGCCGGTCAGTTCGCGGACGC 127

121 gcacgaagacggttcacgtcaagagacagtatggcgttatggcgcgggtcagcgccttcacac 180

Dbb
128 GCAGCAGACGGTGCAGTGAAGGAGAGTGGCGGTCTGGCGGGTCCGCAGTTTCACGAGA 187

181 gcgcgtcagccatgcccacatcaagtcggagatgccatccagcagtcggaggtgttt 240
|||||
188 gcgcgtcagccatgcccacatcaagtcggagatgccatccagcagtcggaggtgttt 247

241 gaagggagcagggaacaaaggtaacctggcagactgttcaagggaagaagggtgtg 300

248 GAAGGGGAGCCAGGGGAACAAGGTGAACCTGGCGAGCTGTTCAAGGGCAAGAAGSGGTG 307

301 ctgattgaggttctctgggccttcaacctggatgttccaagacacacctgccagggttt 360

dbb 308 CTGTTTGGAGTTCCTGGGGCCTTACCCCTGGATGTTCCAAGACACACCTGCCAGGGTTT 367

368 GTGGAGCAGGCTGAGGCTCTGAGGCCAAGGGAGTCCAGGTGGTGGCCCTGCTGAGTGTT 427

421 aatgatgcctttgtgactggcagtgggccgagcccaaggcggaaaggcaaggttcg 480
|||||
428 AATGATGCCCTTTGTGATGGCGAGTGGGCGAGCCCAAGGCGGGAAGSCAAGTTCGG 487
|||||

481 ctcctggctgatcccaactgggcctttgggaagagacagacttattactagatgatcg 540

488 CTCCTGGCTGATCCCACTGGGCGCTTTGGGAAGAGACAGACTATTACTAGATTCG 547

541 ctgggtgccatotttgggaatcgacgtctcaagaggttctccatggtgtacaggatgac 600

548 CTGGTGCCCATCTTTGGGAATCGACGCTCAAGAGGTCTCCATGGTGATACAGGATGEC 607

[illegible]

Db	608	ATATGTGAGG-CCTGAATGTGCAACAGATGCGCAGGCTCAC-----TGCAGCTGGCACC	653
Qy	661	aatatcatctcacagctctgagcgccctggcgccagattactctccacgcctccctatct	720
Db	664	AATATCATCTCACAG-TCTGAGGGCCTGGGCAAGATTACTTCTCCAACCATCC----TAT	718
Qy	721	cacctgcccagccctgtctgctggggccctgcgaattggaatgttggccagatttctgcaata	780
Db	719	TCACTGGACAGGCCTGTGCTGGGGCCCTGCAATTGATTTGGCC--AATTCTGCAATA	775
Qy	781	aacacttgtgtttgcggaaaaaa	805
Db	776	AACACTGTGGGTTTGGGACAACA	800

RESULT 14	
AV650400	
LOCUS	AV650400 677 bp mRNA EST
DEFINITION	AV650400 GLC Homo sapiens cDNA clone GLCCED08 3', mRNA sequence.
ACCESSION	AV650400
VERSION	AV650400.1 GI:9871414
KEYWORDS	EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 677)

AUTHORS

Qian, B., Wu, T., Huang, Q., Huang, C., Kang, B., Gao, X., Xu, Z., Xiao, H., Xu, X., Li, N., Peng, Y., Liu, F., Qu, J., Song, H., Cheng, Z., Qu, J., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gu, Y., Chen, Z., and Han, Z.

TITLE	Homo sapiens cDNA clone
JOURNAL	unpublished (2000)
COMMENT	Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)

Fax: 86-21-50801922
 Email: hanzq@chgc.sh.cn
 This clone is available at CHCC in Shanghai.
 Location/Qualifiers

FEATURES

```

source
1. .677
/organism="Homo sapiens"
/db_xref="taxon:9606"
/allele="ATCATTGG"

```

```
clone_lib="GLC"  
clone_lib="GLC"  
clone_lib="GLC"  
tissue_type="corresponding non cancerous liver tissue"  
dev_stage="Adult"
```

```

/ lab_nost="SOLR"
/ note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
136 a      158 c      248 g      132 t      3 others
BASE COUNT

```

ORIGIN	Query Match	82.9%;	Score 667;	DB 32;	Length 677;
--------	-------------	--------	------------	--------	-------------

```

QY      1  gccaggagccggaatggaaatggccctggagcggaatggaaactaactgactgactgaccc 60
Best Local Similarity 99.7%; Pred. No. 1.2e-162;
Matches 667; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

db 8 GCCAGGAGCGGAGTGGAACTGGCCGTGGGCGGCTATGGGACTAGCTGGCTGTGGCC 67

ov 61 ctgaacagctcagggagctatatactcattcattcagggcgagcgcctcagctctccgcgcgc 120

68 CTGAGACGCTCAGCGGCTATATACTCGCGTGGGGCCGGGGTCACTCTGCGGCAAGC 127

180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998

QY 181 gccctgcagccatggcccaatcaagtgaggatggatgccatccacagcagtgagggtgtt 240
 Db 188 GCCCTGCAGCATGGCCCAATCAAGTGGAGATGGATCCATCCAGCAGTGGTGT 247
 QY 241 gaaggagcaggaagaagtgtaacctgcacagctgttcaaggccaagaagggtg 300
 Db 248 GAAGGGAGCCAGGACAGAGTGAACCTGGCAGAGCTGTTCAGGGCAAGAGGGTGT 307
 QY 301 ctgtttgagttctctggggccttccacccttgatgttccaaagacacacacctccagggtt 360
 Db 308 CTGTTGGAGTCTCTGGGGCTTCCACCTTGATGTTCCAGACACACCTGCCAGGTT 367
 QY 361 gtgagcagctgaggtctgaaggccaaggagtgagtggtggtcgtctgagttt 420
 Db 368 GTGAGCAGCTGAGGCTCTGAAGCGCAAGGGAGTCCAGGTGTGGCTGTCTGAGTGT 427
 QY 421 aatgatcctttgtactggcgagtgggcgagccacacagcggaagcaaggttcg 480
 Db 428 AATGATGCTTGTGACTGGCGAGTGGGGCGAGCCCAAGGGCAAGGCAAGGTTCGG 487
 QY 481 ctctggtgatccactggggccttgggaaggagacagacttattactagattcg 540
 Db 488 CTCCTGCTGATCCACTGGGGCTTGGGAAGGAGACAGACTTATTACTAGATGATCG 547
 QY 541 ctggtgcatctttgggaatogacgtctcaagggttctccatggtgtacagatggc 600
 Db 548 CTGCTGCTCATCTTTGGGAATCGACGCTCAAGAGGTTCTCCATGGTGTACAGGATGGC 607
 QY 601 atagtgaagccctgaatgtggaaccagatggcacagcctcacctcagcctggaacc 660
 Db 608 ATAGTGAAGCCCTGAATGTGGAAACAGATGGCAAGGCTTACCTGCAGNCTGGCACCC 667
 QY 661 aatatcatc 669
 Db 668 AATATCATC 676

RESULT 15

BE796516
 LOCUS 60158985F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3944244 5',
 DEFINITION mRNA sequence.
 ACCESSION BE796516
 VERSION BE796516.1 GI:10217701
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 842)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999).
 Contact: Robert Strausberg, Ph.D.
 Email: cgapb-f@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 Cloning strategy: Incyte Genomics, Inc.
 Cloned through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
 Plate: LCM800 row: m column: 13
 High quality sequence stop: 711.

FEATURES

source
 1..842
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3944244"
 /clone_lib="NIH_MGC_7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:

EcORI; cDNA made by oligo-dT priming. Directionally
 cloned into ECORI/XhoI sites using the following 5'
 adaptor: GGCACAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 193 a 199 c 271 g 179 t
 ORIGIN

Query Match 82.8%; Score 566.4; DB 140; Length 842;
 Best Local Similarity 99.6%; Pred. No. 1.8e-162;
 Matches 689; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 114 ggcagcgcagcaagacagtgacgtgaagagagtgagcgctgctgagcggtccgcagttt 173
 Db 4 GGCAGCGCAGACAGCGTGCAGTGAAGAGAGTGGCGGTCTGGCGGGTCCGAGTTT 63
 QY 174 cagcagagccgtgcagccatggccccaatcaagtgaggatgccatccacagcagtgga 233
 Db 64 CAGCAGAGCCGCTGCAGCCATGGCCCAATCAAGGTGGAGATGCCATCCAGCAGTGA 123
 QY 234 ggtgttgaaggagccaggggaacaagtgtaacctggcagagctgttcaaggggcaagaa 293
 Db 124 GGTGTTGAAGGGAGGCA-GGAACAGGTGAACCTGGCAGAGCTGTTCAGAGGCAAGAA 182
 QY 294 ggggtgctgtttgagttcctgggcttcaccctgagtggttccaaagacacacctcc 353
 Db 183 GGGTGTGCTGTTTGGAGTCTCTGGGGCTTCACCCCTGGATGTTCACAGACACACTGCC 242
 QY 354 aggtttgtgagcaggtcctgaggtctgaagccaaaggagtcaggtggtggcctgtct 413
 Db 243 AGGTTTGTGGAGCAGCGCTGAGGCTCTGAGCCCAAGGAGTCCAGGTGGTGGCTGCT 302
 QY 414 gagtgttaataatgcctttgtgactggcgagtgggggccgagccacaagggcggaagcaa 473
 Db 303 GAGTGTTAATGATGCTCTTGTGACTGGCGAGTGGGGCGGAGCCACAAAGGGGAAGCAA 362
 QY 474 gttcgcctcctgctaatccactgggcttgggaaggagacagacttattactaga 533
 Db 363 GGTTCGGCTCTGGCTGATCCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGA 422
 QY 534 tgattcgtgtgtccatctttgggaatcgacgtctcagaggttctccatggtgttaca 593
 Db 423 TGATTGCTGTGTCCTCTTTGGGAATCGAGTCTCAAGAGTCTCCATCGTGTGTACA 482
 QY 594 ggatggcatagtgaaggccctgtaattgtgaaaccagatggcacagcctcactcagcct 653
 Db 483 GGATGGCATAGTGAAGGCCCTGAATGTGGAACAGATGGCAGCCCTCACCTGCAGCT 542
 QY 654 ggcacccaataatcctcaagctctgagccctgggagcagattactcctccaccctc 713
 Db 543 GGCACCAATAATCATCTCACAGCTCTGAGGCCCTGGGGCAGATTAATCTCCCTCCACCCCTC 602
 QY 714 cctatctcactgcacagcctgtgctggtggccctgcgaatggaaatgttggccagatttc 773
 Db 603 CCTATCTCACCTGCCAGCCCTGTGCTGGGGCCCTGCAATTTGGAATGTTGG-CAGATTTC 661
 QY 774 tgcataaacaactgtgtgttgcggaaaaaa 805
 Db 662 TGCAATAAACACTTGTGTTGTTGGCGCAAAAAA 693

Search completed: October 22, 2001, 11:50:25
 Job time: 4140 sec

OM of US-09-486-167A-2 to: EST:* out_format : pfs

Date: Oct 22, 2001 10:19 AM

About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framer_p2n.model -DEV=xlp
-Q/cn2_1/USPTO_spool/US09486167/runat_22102001_064206_25012/app_query.fasta_1.222
-DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-DELOP=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-LIST=7.000 -START=1 -MATRIX=biosum62 -TRANS=human40.cdi
-ALIGN=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USER=US09486167 @CGNL_1.2970 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-486-167A-2

Query length: 162

Database: EST*

Database sequences: 10228115

Database length: 431459454

Search time (sec): 1294.880000

score_list:

Sequence	Strd	Orig	ZScore	Escore	Len	Documentation
gb_est72:BE298055	+	830.00	1740.13	8.3e-88	604	! BE298055 601118369F1 NIH_MGC_17
gb_est98:BG355577	+	830.00	1739.83	8.6e-88	622	! BE298055 602368257F1 NIH_MGC_91
gb_est95:BF971772	+	830.00	1739.81	8.6e-88	623	! BF971772 602240049F1 NIH_MGC_46
gb_est98:BG287024	+	830.00	1738.55	1.0e-87	703	! BG287024 602382319F1 NIH_MGC_93
gb_est99:BG325661	+	830.00	1738.45	1.0e-87	710	! BG325661 602424378F1 NIH_MGC_14
gb_est74:BE388478	+	830.00	1738.21	1.1e-87	726	! BE388478 601287129F1 NIH_MGC_44
gb_est29:AL541916	+	830.00	1735.76	1.5e-87	918	! AL541916 601306118F1 NIH_MGC_3
gb_est78:BE736544	+	830.00	1734.57	1.7e-87	1029	! BE736544 601306118F1 NIH_MGC_3
gb_est46:AW09904	+	826.00	1732.37	2.2e-87	565	! AW09904 601306118F1 NIH_MGC_3
gb_est40:AV650400	+	813.00	1702.98	9.7e-86	677	! AV650400 601306118F1 NIH_MGC_3
gb_est29:AL535869	+	810.00	1694.66	2.8e-85	817	! AL535869 601306118F1 NIH_MGC_3
gb_est79:BE784530	+	802.00	1678.10	2.4e-84	790	! BE784530 601474133F1 NIH_MGC_68
gb_est72:BE300059	+	802.00	1675.55	3.3e-84	1009	! BE300059 600944516F1 NIH_MGC_1
gb_est80:BE893687	+	801.00	1677.78	2.5e-84	665	! BE893687 601436488F1 NIH_MGC_72
gb_est92:BE792903	+	801.00	1677.01	2.7e-84	716	! BE792903 602253206F1 NIH_MGC_84
gb_est72:BE300398	+	801.00	1676.51	2.9e-84	751	! BE300398 600944516F1 NIH_MGC_17
gb_est91:BF669149	+	801.00	1676.01	3.1e-84	788	! BF669149 602120006F1 NIH_MGC_56
gb_est79:BE795522	+	801.00	1675.39	3.3e-84	836	! BE795522 601522505F1 NIH_MGC_7
gb_est79:BE796516	+	801.00	1675.32	3.4e-84	842	! BE796516 601589855F1 NIH_MGC_7
gb_est92:BF794134	+	801.00	1675.31	3.4e-84	843	! BF794134 602255506F1 NIH_MGC_85
gb_est76:BE562997	+	801.00	1675.28	3.4e-84	845	! BE562997 601336252F1 NIH_MGC_44
gb_est92:BF791499	+	799.00	1668.48	3.6e-84	884	! BF791499 601305314F1 NIH_MGC_38
gb_est28:AL541915	+	796.00	1666.17	1.1e-83	1081	! AL541915 602255506F1 NIH_MGC_85
gb_est79:BE781787	+	796.00	1665.85	1.1e-83	734	! BE781787 601470589F1 NIH_MGC_67
gb_est21:BG498369	+	796.00	1664.88	1.3e-83	831	! BG498369 60244160F1 NIH_MGC_60
gb_est29:AL558438	+	795.00	1665.64	1.2e-83	631	! AL558438 601300587F1 NIH_MGC_21
gb_est28:AL558439	+	795.00	1664.13	1.4e-83	729	! AL558439 601300587F1 NIH_MGC_21
gb_est28:AL520417	+	794.00	1661.84	1.9e-83	741	! AL520417 601305314F1 NIH_MGC_38
gb_est78:BE729013	+	792.00	1657.09	3.5e-83	779	! BE729013 601562279F1 NIH_MGC_20
gb_est101:BG482283	+	790.00	1653.12	5.8e-83	760	! BG482283 602526614F1 NIH_MGC_21
gb_est97:BG113765	+	790.00	1652.66	6.2e-83	794	! BG113765 602284293F1 NIH_MGC_86
gb_est74:BE410064	+	789.00	1652.57	5.7e-83	611	! BE410064 601300587F1 NIH_MGC_21
gb_est95:BF976165	+	786.00	1648.57	1.0e-82	959	! BF976165 602245026F1 NIH_MGC_4
gb_est28:AL520418	+	783.00	1637.14	4.5e-82	850	! AL520418 602245026F1 NIH_MGC_4
gb_est95:BF975906	+	781.00	1632.44	8.3e-82	889	! BF975906 602246103F1 NIH_MGC_48
gb_est40:AV705034	+	777.00	1626.37	1.8e-81	707	! AV705034 602573237F1 NIH_MGC_77
gb_est102:BG545942	+	770.00	1609.95	1.5e-80	825	! BG545942 602573237F1 NIH_MGC_77
gb_est78:BE745244	+	768.50	1607.53	2.0e-80	768	! BE745244 601574011F1 NIH_MGC_9
gb_est16:AT119223	+	767.00	1608.81	1.7e-80	501	! AT119223 ue94h08.y1 Sugano mous

gb_est13:AA869971 + 767.00 1607.90 1.9e-80 547 ! AA869971 vql9h12.r1 Barstead
gb_hic:AK002383 + 767.00 1604.38 3.0e-80 766 ! AK002383 Mus musculus adult
gb_hic:AK003332 + 767.00 1604.38 3.0e-80 766 ! AK003332 Mus musculus 18 day
gb_est100:BG403357 + 766.50 1601.53 4.4e-80 910 ! BG403357 602419031F1 NIH_MGC_1

seq_name: gb_est72:BE298055

seq_documentation_block:
LOCUS BE298055 604 bp mRNA EST
DEFINITION 601118369F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3028312 5',
mRNA sequence.

ACCESSION BE298055
VERSION BE298055.1 GI:9181641
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 604)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LCM86 row: i column: 17
High quality sequence stop: 601.
FEATURES
Location/Qualifiers
1..604

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3028312"
/clone_lib="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: muscle; Vector: pOTB7; Site: 1: EcoRI;
Site: 2: XhoI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 123 a 158 c 194 g 129 t
ORIGIN
alignment_scores:
Quality: 830.00 Length: 162
Ratio: 5.123 Gaps: 0
Percent Similarity: 100.00 Percent Identity: 100.000
alignment_block:
US-09-486-167A-2 x BE298055 ..
Align seg 1/1 to: BE298055 from: 1 to: 604

1 MetalaProfileValGlyAspAlaIleProAlaValGluValPheG1 17
|||||
38 ATGGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGTGTGA 87
|||||

17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
88 AGGGAGCCAGGAACAGGTGAACCTGGCAGAGCTGTTCAGGCAAGA 137
|||||

34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
138 AGGGTGTGTCTTTGGAGTTCTGGGGCTTCACCCCTGGATGTTCAAG 187
|||||

alignment_scores: Quality: 830.00 Length: 162
Ratio: 5.123 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-486-167A-2 x BG255777 ..
Align seg 1/1 to: BG255777 from: 1 to: 622
1 MetAlaProIleLeuValGlyAspAlaIleProAlaValGluValPheG1 17
109 ATGGCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTGA 158
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34
159 AGGGAGCCAGGGAACAAGTGAACCTGGCAGAGCTGTCAAGGCCAAGA 208
34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
209 AGGTGTGCTGTTTGGAGTTCTCTGGGGCTTCCACCTGGATGTTCCAAG 258
51 ThrHisLeuProGlyPheValGluInAlaGluAlaLeuLysAlaLysG1 67
259 ACACACCTGCCAGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 308
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
309 AGTCCAGGTGGTGGCTGCTGAGTGTAAATGATGCTTTGTGACTGGCG 358
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
359 AGTGGGCCCGACGACCAAGCGGAGGAGGTTCCGCTCTCGCTGCTGAT 408
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
409 CCCACTGGGCTTTGGAGGAGCAGACTTATTACTAGATGATTCGCT 458
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
459 GGTGTCCATCTTTGGGAATCAGCTCTCAGAGGTTCTCCATGGTGGTAC 508
134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
509 AGGATGCGATAGTGAAGCCCTGAATGTGGAAACAGATGGCAGGCGCTC 558
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
559 ACCTGCAGCTGGCACCACCAATATCATCTCACAGCTC 594
seq_name: gb_est95:BF971772
seq_documentation_block:
LOCUS BF971772 622 bp mRNA EST 22-JAN-2001
DEFINITION G02240049F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4328364 5',
mRNA sequence.
ACCESSION BF971772
VERSION BF971772.1 GI:12338987
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 622)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
Tissue Procurement: ATCC
cdNA Library Preparation: Ling Hong/Rubin Laboratory
cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
188 ACACACCTGCCAGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 237
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
238 AGTCCAGGTGGTGGCTGCTGAGTGTAAATGATGCTTTGTGACTGGCG 287
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
288 AGTGGGCCGAGCCACACAGGCGGAAGCAAGTTCGGCTCTCGCTGAT 337
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
338 CCCACTGGGCTTTGGAGGAGGAGACAGACTTATTACTAGATGATTCGCT 387
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
388 GGTGTCACTTTTGGGAATCAGCTCTCAGAGGTTCTCCATGGTGGTAC 437
134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
438 AGGATGCGATAGTGAAGCCCTCAATGTGGAACAGATGGCAGGCGCTC 487
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
488 ACCTGCAGCTGGCACCACCAATATCATCTCACAGCTC 523
seq_name: gb_est98:BG255777
seq_documentation_block:
LOCUS BG255777 622 bp mRNA EST 13-FEB-2001
DEFINITION G02368257F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4476207 5',
mRNA sequence.
ACCESSION BG255777
VERSION BG255777.1 GI:12765593
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 622)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
Tissue Procurement: DCTD/DTP
cdNA Library Preparation: Life Technologies, Inc.
cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10303 row: b column: 16
High quality sequence stop: 622.
Location/Qualifiers
1..622
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_91"
/issue_type="adenocarcinoma, cell line"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally; oligo-AT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
126 a 153 c 216 g 127 t
BASE COUNT
ORIGIN

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM1188 row: b column: 13
 High quality sequence stop: 619.

FEATURES

source
 1. 623
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4328364"
 /clone_lib="NIH_MGC_46"
 /tissue_type="leiomyosarcoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: uterus; Vector: pOTB7; Site:1: XhoI; Site:2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCCAGGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."
 BASE COUNT 127 a 157 c 205 g 133 t 1 others
 ORIGIN

alignment_scores:
 Quality: 830.00 Length: 162
 Ratio: 5.123 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-486-167A-2 x BF971772 ..
 Align seg 1/1 to: BF971772 from: 1 to: 623

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
 |||||
 56 ATGGCCCAATCAAGTGGAGATGCCATCCAGCAGTGGAGGTGTTGA 105
 |||||
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysL 34
 |||||
 106 AGGGAGCCAGGACAAAGTGAACCTGCAGAGCTGTCAAGGGCAAGA 155
 |||||
 34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
 |||||
 156 AGGGTGTGCTGTTGGAGTCTCTGGGGCTTCACCCCTGGATGTCCAAG 205
 |||||
 51 ThrHisLeuProGlyPheValGluGlnAlaLeuLysAlaLysG1 67
 |||||
 206 ACACACCTGCCAGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 255
 |||||
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 |||||
 256 AGTCAGGTGGTGGCTGTCTGAGTGTAAATGATGCCCTTTGTGACTGGCG 305
 |||||
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
 |||||
 306 AGTGGGGCCGACCCACAAAGCGGAAGCAAGGTTCCGCTCTCGCTGAT 355
 |||||
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
 |||||
 356 CCCATCTGGGGCTTTGGGAAGGAGACAGACTATTACTAGATGATTCGCT 405
 |||||
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
 |||||
 406 GGTGTCCATCTTTGGGAATGCACGCTCTCAAGAGGTTCTCCATGGTGGTAC 455
 |||||
 134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
 |||||
 456 AGGATGGCATAGTGAAGCCCTGAATGTGGAAACCAGATGGCACAGGCCCTC 505
 |||||

151 ThrCysSerLeuAlaProAsnIleSerGlnLeu 162
 |||||
 506 ACCTGCAGCCCTGGCACCACCATATCATCTACAGCTC 541
 |||||

seq_name: gb_est98:BG287024

seq_documentation_block:
 LOCUS BG287024 703 bp mRNA EST 21-FEB-2001
 DEFINITION 602382319F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4499716 5',
 mRNA sequence.

ACCESSION BG287024

VERSION BG287024

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 703)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10364 row: f column: 05

High quality sequence stop: 702.

FEATURES

Location/Qualifiers

Source

1. 703

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4499716"

/clone_lib="NIH_MGC_93"

/tissue_type="transitional cell papilloma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: bladder; Vector: pCMV-SPORT6; Site:1: NotI;

Site:2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.7 Kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 142 a 177 c 234 g 150 t

ORIGIN

alignment_scores:

Quality: 830.00 Length: 162
 Ratio: 5.123 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-486-167A-2 x BG287024 ..

Align seg 1/1 to: BG287024 from: 1 to: 703

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17

|||||

105 ATGCCCAATCAAGTGGAGATGCCATCCAGCAGTGGAGGTGTTGA 154

|||||

17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysL 34

|||||

155 AGGGAGCCAGGGAACAAAGTGAACCTGGCAGAGCTGTCAAGGGCAAGA 204

|||||

34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50

|||||

205 AGGGTGTGCTTTGGAGTTCCTGGGGCTTCACCCCTGGATGTCCAAG 254

|||||

51 ThrHisLeuProGlyPheValGluGlnAlaLeuLysAlaLysG1 67

|||||

255 ACACACCTCCAGGGTTTGTGGAGCAGGCTGTGAGGCTCTGAAGGCCAAGGG 304
 67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 |||||
 305 AGTCCAGGTGGTGGCTGTCTGAGTGTAAATGATGCTTTGTGACTGGCG 354
 |||||
 84 luTrpGlyArgAlaHisTysAlaGluGlyLysValArgLeuLeuAlaAsp 100
 |||||
 355 AGTGGGGCCGAGCCCAAGGCGGAAGGCAAGGTTCCGGCTCCTGGCTGAT 404
 |||||
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuLeuAspSerLe 117
 |||||
 405 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 454
 |||||
 117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
 |||||
 455 GGTGTTCATCTTTGGGAATCGAGTCTCAAGAGGTTCCTCCATGGTGTAC 504
 |||||
 134 luAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
 |||||
 505 AGGATGGCATAGTGAAGGCCCTGAATGTGAACCAAGATGGCACAGGCCTC 554
 |||||
 151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
 |||||
 555 ACCTGCAGCTGGCACCAATATCATCTCACAGCTC 590
 |||||

seq_name: gb_est99:BG325661

seq_documentation_block: 710 bp mRNA EST 27-FEB-2001
 LOCUS BG325661 602424378F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562443 5',
 DEFINITION mRNA sequence.
 ACCESSION BG325661 GI:13132098
 VERSION BG325661
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 710)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/BTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L10M1274 row: k column: 20
 High quality sequence stop: 678.
 Location/Qualifiers
 1..710
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4562443"
 /clone_lib="NIH_MGC_14"
 /tissue_type="renal cell adenocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5',
 adaptor: GGACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

174 a 167 c 222 g 144 t

FEATURES

source

seq_name: gb_est74:BE388478
 seq_documentation_block: 726 bp mRNA EST 21-JUL-2000
 LOCUS BE388478 601287129F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3621583 5',
 DEFINITION mRNA sequence.
 ACCESSION BE388478
 VERSION BE388478.1 GI:9333843
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 726)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

alignment_scores:

Quality: 830.00 Length: 162
 Ratio: 5.123 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-486-167A-2 x BG325661 ..

Align seg 1/1 to: BG325661 from: 1 to: 710

1 MetAlaProIleLysValGlyAspAlaIleProAlaValAluValPheG 17
 |||||
 38 ATGGCCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 87
 |||||
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLys 34
 |||||
 88 AGGGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 137
 |||||
 34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
 |||||
 138 AGGGTGTGCTGTTGGAGTTCCTGGGGCTTCACCCCTGGATGTTCCAAG 187
 |||||
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67
 |||||
 188 ACACACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 237
 |||||
 67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 |||||
 238 AGTCCAGGTGGTGGCTGTCTGAGTGTAAATGATGCTTTGTGACTGGCG 287
 |||||
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
 |||||
 288 AGTGGGGCCGAGCCCAAGGCGGAAGGCAAGGTTCCGGCTCCTGGCTGAT 337
 |||||
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
 |||||
 338 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 387
 |||||
 117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
 |||||
 388 GGTGTTCATCTTTGGGAATCGAGTCTCAAGAGGTTCCTCCATGGTGTAC 437
 |||||
 134 luAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
 |||||
 438 AGGATGGCATAGTGAAGGCCCTGAATGTGAACCAAGATGGCACAGGCCTC 487
 |||||
 151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
 |||||
 488 ACCTGCAGCTGGCACCAATATCATCTCACAGCTC 523
 |||||

DNA Sequencing by: Inocyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: L1CM295 row: 1 column: 08
 High quality sequence stop: 638.

FEATURES

source
 1. .726
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3621583"
 /clone_lib="NH_MGC_44"
 /cissue_type="endometrium, adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCAGGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 161 a 171 c 233 g 161 t
 ORIGIN

alignment_scores:
 Quality: 830.00 Length: 162
 Ratio: 5.123 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-486-167A-2 x BE388478 ..
 Align seg 1/1 to: BE388478 from: 1 to: 726

1 MetAlaProLeLysValGlyAspAlaIleProAlaValGluValPheG1 17
 |||||
 50 ATGGCCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTTGA 99
 |||||
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
 |||||
 100 AGGGAGCCAGGGAACAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 149
 |||||
 34 ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
 |||||
 150 AGGTGTGTCTGTTGGAGTTCCTGGGGCTTCACCCCTGGATGTTCCAAG 199
 |||||
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysGly 67
 |||||
 200 ACACCTGCCAGGGTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 249
 |||||
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 |||||
 250 AGTCCAGGTGGTGTCTGAGTGTAAATGATGCTTTGTGACTGGCG 299
 |||||
 84 lutrPGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
 |||||
 300 AGTGGGCCAGCCACAGCGGAAAGGTTCCGGCTCTCGGCTGAT 349
 |||||
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117
 |||||
 350 CCCACTGGGGCTTTGGGAGGAGACAGACTTATTACTAGATGATTCGCT 399
 |||||
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
 |||||
 400 GGTGTCCATCTTGGGAATCGAGCTCTCAAGAGTTCTCCATGGTGTAC 449
 |||||
 134 lnaSPGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
 |||||
 450 AGGATGGCATAGTGAAGGGCCCTGAATGTGGAAACAGATGGCAGGCCCTC 499
 |||||
 151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
 |||||

500 ACCTGCAGCTGGCACCACCAATATCATCTCACAGCTC 535

seq_name: gb_est29:AL541916

seq_documentation_block:

LOCUS AL541916 918 bp mRNA EST 16-FEB-2001
 DEFINITION AL541916 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE007YL04 5 prime
 , mRNA sequence.

ACCESSION AL541916

VERSION AL541916.1 GI:12873445

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 918)

AUTHORS Li,W.B., Gruber,C., Jesse,J., Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. .918
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DE007YL04"
 /clone_lib="LTI_FL002_PL1"
 /lab_host="DH10B"
 /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
 cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-stranded cDNA was digested with Not I
 and cloned into the Not I and Eco RV sites of the
 pCMVSPORT 6 vector. Library was constructed by Life
 Technologies. Contact : feng liang Life Technologies, a
 division of Invitrogen 9800 Medical Center Drive Rockville
 , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
 fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 218 a 225 c 296 g 178 t
 ORIGIN

alignment_scores:
 Quality: 830.00 Length: 162
 Ratio: 5.123 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-486-167A-2 x AL541916 ..

Align seg 1/1 to: AL541916 from: 1 to: 918

1 MetAlaProLeLysValGlyAspAlaIleProAlaValGluValPheG1 17
 |||||
 254 ATGGCCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTTGA 303
 |||||
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
 |||||
 304 AGGGAGCCAGGGAACAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 353
 |||||
 34 ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
 |||||
 354 AGGTGTGTCTGTTGGAGTTCCTGGGGCTTCACCCCTGGATGTTCCAAG 403
 |||||
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysGly 67
 |||||
 404 ACACCTGCCAGGGTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 453
 |||||
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 |||||
 454 AGTTCAGGTGGTGGCTCTCTGAGTGTAAATGATGCCTTTGTGACTGGCG 503
 |||||

```

84  luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAasp 100
|||||
504  AGTGGGCGCCAGCCACAGGCGGAAGGTCGGCTCGCTGGCTGAT 553
|||||
101  ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAaspSerLe 117
|||||
554  CCCACTGGGGCCCTTGGGAAGGAGAGACAGACTTATTACTAGATCGCT 603
|||||
117  uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
|||||
604  GGTGTCCATCTTGGGAATCGAGCTCAAGAGGTCTCCATGGTGGTAC 653
|||||
134  lnaSpGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
654  AGGATGGCATAGTGAAGGCCCTGAATGTGAACACAGATGGCAGGCGCTC 703
|||||
151  ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
|||||
704  AACTGAGCGCTGGCACCACCAATATCATCTCACAGCTC 739
|||||

```

seq_name: gb_est78:BE736544

```

seq_documentation_block:
LOCUS BE736544 1029 bp mRNA EST 15-SEP-2000
DEFINITION 601306118F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3640276 5',
mRNA sequence.
ACCESSION BE736544
VERSION BE736544.1 GI:10150536
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM344 row: d column: 05
High quality sequence stop: 658.
Location/Qualifiers
1..1029
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3640276"
/clone_lib="NIH_MGC_39"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: pOTB7; Site: 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."
```

BASE COUNT
ORIGIN

299 a 234 c 321 g 175 t

alignment_scores:

Quality: 830.00 Length: 162
Ratio: 5.123 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-486-167A-2 x BE736544

Align seg 1/1 to: BE736544 from: 1 to: 1029

```

1  MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheGl 17
|||||
51  ATGGCCCCAATCAAGGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 100
|||||
17  uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
101  AGGGGAGCCAGGGAACCTGACCTGGCAGAGCTGTTCAAGGGCAAGA 150
|||||
34  ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
151  AGGGTGTGCTGTTGGAGTTCCTGGGGCCCTCACCCCTGGATGTTCCAAG 200
|||||
51  ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysGl 67
|||||
201  ACACACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 250
|||||
67  yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
251  AGTCCAGGTGGTGGCCCTGTCTGAGTGTAAATGATGCCCTTTGTGACTGG 300
|||||
84  luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAasp 100
|||||
301  AGTGGGCGCGAGCCACAGCGGGAAGGCAAGGTTGCGCTCCTGGCTGAT 350
|||||
101  ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAaspSerLe 117
|||||
351  CCCACTGGGGCCCTTGGGAAGGAGACAGACTTATTACTAGATCATTCGTT 400
|||||
117  uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
|||||
401  GGTGTCCATCTTGGGAATCGAGCTCTCAAGAGGTTCCTCATGGTGGTAC 450
|||||
134  lnaSpGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
451  AGGATGGCATAGTGAAGGCCCTGAATGTGAACACAGATGGCAGGCGCTC 500
|||||
151  ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
|||||
501  AACTGAGCGCTGGCACCACCAATATCATCTCACAGCTC 536
|||||

```

seq_name: gb_est46:AW409904

```

seq_documentation_block:
LOCUS AW409904 565 bp mRNA EST 29-JUN-2000
DEFINITION fh03a11.x1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960948 5',
mRNA sequence.
ACCESSION AW409904
VERSION AW409904.1 GI:6935445
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 565)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Plate: L1CM55, row: B, column: 21
 Seq primer: -21M13 forward primer (ABI).
 Location/Qualifiers

FEATURES

Source
 1..565
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2960948"
 /clone_lib="NIH_MGC_17"
 /tissue_type="rhabdomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: muscle; Vector: pORF7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 119 a 133 c 197 g 116 t
 ORIGIN

alignment_scores:
 Quality: 826.00 Length: 161
 Ratio: 5.130 Gaps: 0
 Percent Similarity: 100.000 Percent identity: 100.000

alignment_block:

US-09-486-167A-2 x AW409904 ..

Align seg 1/1 to: AW409904 from: 1 to: 565

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
 82 ATGGCCCCAATCAAGTGGAGATCCATCCAGCAGTGGAGGTGTGA 131
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysL 34
 132 AGGGAGCCAGGAACAAGTGAACCTGGCAGAGCTTCAAGGCCAAGA 181
 34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
 182 AGGGTGTGCTTTGGAGTCTCTGGGGCTTACCCCTGGATGTCCAAG 231
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
 232 ACACACCTGCCAGGTTTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 281
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 282 AGTCCAGGTGGTGGCTGTCTGAGTGTAAATGATGCCCTTTGTGACTGGCG 331
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
 332 AGTGGGGCCGAGCCACAAAGCGGAAGGTCAGGCTCTGCTGCTGAT 381
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
 382 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATCGCT 431
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
 432 GGTGTCCATCTTTGGGAATGACGCTCTCAAGAGGTTCCTCCATGGTGTAC 481
 134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
 482 AGGATGGCATAGTAAGGCCCTGAATGGAAACAGGATGGACAGGCGCTC 531
 151 ThrCysSerLeuAlaProAsnIleIleSerGln 161
 532 ACCTGACGCTGGCACCACCAATATCATCTCACAG 564

seq_name: gb_est40:AV650400

seq_documentation_block:

LOCUS AV650400 677 bp mRNA EST 07-SEP-2000
 DEFINITION AV650400 GLC Homo sapiens cDNA clone GLCCED08 3', mRNA sequence.
 ACCESSION AV650400
 VERSION AV650400.1 GI:9871414
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 677)
 AUTHORS Qian,B., Wu,T., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H.,
 Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Qu,J.,
 Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
 ,G., Yang,Y., Gu,Y., Chen,Z. and Han,Z.
 Homo sapiens cDNA clone
 Unpublished (2000)
 COMMENT
 Contact: Zequang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.

FEATURES

source

1..677
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="GLCCED08"
 /clone_lib="GLC"
 /tissue_type="corresponding non cancerous liver tissue"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 136 a 158 c 248 g 132 t 3 others
 ORIGIN

alignment_scores:

Quality: 813.00 Length: 159
 Ratio: 5.113 Gaps: 0
 Percent Similarity: 100.000 Percent identity: 99.371

alignment_block:

US-09-486-167A-2 x AV650400 ..

Align seg 1/1 to: AV650400 from: 1 to: 677

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
 200 ATGGCCCCAATCAAGTGGAGATGCCATCCAGCAGTGGAGGTGTGA 249
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysLysL 34
 250 AGGGAGCCAGGACAAAGTGAACCTGGCAGAGCTTCAAGGCCAAGA 299
 34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
 300 AGGGTGTGCTTTGGAGTCTCTGGGGCTTACCCCTGGATGTCCAAG 349
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
 350 ACACACCTGCCAGGTTTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 399
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 400 AGTCCAGGTGGTGGCTGTCTGAGTGTAAATGATGCCCTTTGTGACTGGCG 449
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
 |||||||||||||||||||||||||||||||||||||||||||||||||||

```

450 AGTGGGCGGAGCCCAAGGCGGAGGAGGTTCCGGCTCTCGGCTGAT 499
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117
|||||
500 CCCACTGGGCGCTTTGGGAGGAGACAGACTTATTACTAGATCGCT 549
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetVal 134
|||||
550 GGTGTCATCTTTGGGAATCGAGCTCTCAAGAGGTTCTCCATGGTG 599
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
600 AGGATGGCATAGTGAAGGCGCTGAATGTGGAACCAAGATGGCAGAGCC 649
151 ThrCysSerLeuAlaProAsnIleLe 159
|||||
650 ACCTGCAGNCTGGCACCCCAATATCATC 676

seq_name: gb_est29:AL535869
seq_documentation_block:
LOCUS AL535869 818 bp mRNA EST 13-FEB-2001
DEFINITION AL535869 LTI_FL013.FBrn1 Homo sapiens cDNA clone CS0DF013YH13 5
prime.mRNA sequence.
ACCESSION AL535869
VERSION AL535869.1 GI:12799362
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 818)
AUTHORS Li.W.B., Gruber.C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..818
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DF013YH13"
/dev_stage="pooled tissue from post conception fetuses (20
week, 24 week and 26 week)"
/lab_host="DH10B"
/note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
Technologies. Contact: Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@life.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 176 a 201 c 274 g 165 t 2 others
ORIGIN

```

```

alignment_scores:
  Quality: 810.00 Length: 162
  Ratio: 5.031 Gaps: 1
Percent Similarity: 99.383 Percent Identity: 98.765

```

```

alignment_block:
US-09-486-167a-2 x AL535869 ..

```

```

Align seg 1/1 to: AL535869 from: 1 to: 818

```

```

2 AlaProIleLysValGlyAspAlaIleProAlaValGluValPheGluG1 18

```

```

|||||
190 GCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTTGAAGG 239
18 yGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysLysG 35
|||||
240 GGAGCCAGGGAACAAGTGAACCTGGCAGAGCTGTTCAAGGGCAAGAGG 289
35 lyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLysThr 51
|||||
290 GTGTGCTGTTTGGAGTTCTCTGGGGCCTTCACCCCTGGATGTTCCAA 339
52 HisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysGlyVa 68
|||||
340 CACCTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCCAAGGAG 389
68 lGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyGlu 85
|||||
390 CCAGGTGGTGGCTGCTGAGTGTTAATGATGCTTTGTGACTGGCGAGT 439
85 rpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAspPro 101
|||||
440 GGGGCCGAGCCCAAGGCGGAAGCAAGGTTCCGCTCCTGGCTGATCCC 489
102 ThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLeuVa 118
|||||
490 ACTGGGGCCTTTGGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 539
118 lSer.IlePheGlyAsnArgArgLeuLysArgPheSerMetValValGln 134
|||||
540 GTCCACTCTTTGGGATCCACGCTCTCAAGAGGTTCTCCATGGGTACAG 589
135 AspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeuTh 151
|||||
590 GATGGCATAGTGAAGGCCCTGAATGTGGAACCAAGATGGCACAGGCCTC 639
151 rCysSerLeuAlaProAsnIleLeSerGlnLeu 162
|||||
640 CTGCAGCCTGGCACCCCAATATCATCTCACAGCTC 673

seq_name: gb_est79:BE784530
seq_documentation_block:
LOCUS BE784530 790 bp mRNA EST 20-OCT-2000
DEFINITION 601474153P1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3876972 5',
mRNA sequence.
ACCESSION BE784530
VERSION BE784530.1 GI:10205815
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 790)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsbbs-re@mail.nih.gov
Tissue Procurement: DCTD/DPH/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9638 row: j column: 13
High quality sequence stop: 733.
Location/Qualifiers
1..790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3876972"
/clone_lib="NIH_MGC_68"

```

```

FEATURES
Source

```

/tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
 Average insert size 1.8 kb. Library constructed by Life
 Technologies."
 BASE COUNT 161 a 193 c 273 g 163 t
 ORIGIN

alignment_scores:
 Quality: 802.00 Length: 162
 Ratio: 4.951 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-486-167A-2 x BE784530 ..
 Align seg 1/1 to: BE784530 from: 1 to: 790

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
 |||||||
 182 ATGGCCCCAATCAAGGTGGGAGATGCCATCCAGCAGTGGAGTGTGGA 231
 |||||||
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysL 34
 |||||||
 232 AGGGAGCCAGGACAGTGAAGTGGCAGAGCTGTCAAGGGCAAGA 281
 |||||||
 34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
 |||||||
 282 AGGGTGTGCTTTGGAGTTCCTGGGGCTTCACCCCTGGATGTTCCAA 331
 |||||||
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysG1 67
 |||||||
 332 ACACACCTGCCAGGCTTTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 381
 |||||||
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 |||||||
 382 AGTCCAGGTGGTGGCTGTCTGAGTGTAAATGATGCCCTTTGTACTGGC 431
 |||||||
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
 |||||||
 432 AGTGGGGCCGAGCCACAAAGCGGAAGCAAGGTTGGCTCTCTGGCTGAT 481
 |||||||
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
 |||||||
 482 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATCGCT 531
 |||||||
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
 |||||||
 532 GGTGTCCATCTTTGGATCCACGCTCTCAAGAGTCTCTCCATGGTGTAC 581
 |||||||
 134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
 |||||||
 582 AGGATGGCATAGTGAAGGCCCTGAATGTGAACACAGATGGCACAGGCTC 631
 |||||||
 151 ThrCysSerLeuAlaProAsnIleSerGlnLeu 162
 |||||||
 632 AC.TGCACCTGGCACCCCAATATCATCTCACAGCTC 666

seq_name: gb_est72:BE300059

seq_documentation_block:
 LOCUS BE300059 1009 bp mRNA EST 20-JUL-2000
 DEFINITION 600944516T1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960859 3',
 mRNA sequence.
 ACCESSION BE300059
 VERSION BE300059.1 GI:9183807
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1009)
 NIH-MGC http://mgc.nci.nih.gov/
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCW54 row: 0 column: 04
 High quality sequence start: 20
 High quality sequence stop: 724.

FEATURES

source
 1..1009
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2960859"
 /clone_lib="NIH_MGC_17"
 /lab_host="DH10B (phage-resistant)"
 /tissue_type="rhabdomyosarcoma"
 /note="Organ: muscle; Vector: pOTB7; Site: 1: EcoRI;
 Site 2: XhoI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 239 a 307 c 247 g 216 t
 ORIGIN

alignment_scores:

Quality: 802.00 Length: 162
 Ratio: 4.951 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-486-167A-2 x BE300059/rev ..

Align seg 1/1 to reverse of: BE300059 from: 1 to: 1009

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
 |||||||
 555 ATGGCCCCAATCAAGGTGGGAGATGCCATCCAGCAGTGGAGTGTGGA 506
 |||||||
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysL 34
 |||||||
 505 AGGGAGCCAGGACAAAGTGAACCTGGCAGAGCTGTCAAGGGCAAGA 456
 |||||||
 34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
 |||||||
 455 AGGGTGTGCTTTGGAGTTCCTGGGGCTTCAC.CCTGGATGTTCCAA 407
 |||||||
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysG1 67
 |||||||
 406 ACACACCTGCCAGGCTTTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 357
 |||||||
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 |||||||
 356 AGTCCAGGTGGTGGCTGTCTGAGTGTAAATGATGCCCTTTGTACTGGC 307
 |||||||
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
 |||||||
 306 AGTGGGGCCGAGCCACAAAGCGGAAGCAAGGTTGGCTCTCTGGCTGAT 257
 |||||||
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
 |||||||
 256 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATCGCT 207

117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
|||||
206 GGTGTCCATCTTTGGGAATCGACGCTCAAGAGGTCTCCATGGTGGTAC 157
|||||
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
156 AGATGGCATAGTAGAGGCCCTGGAATGTGGACCATGGTGGACAGGCCCTC 107
|||||
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
|||||
106 ACCTGCAGCGCTGGCACCACCAATATCATCTCACAGCTC 71
|||||
seq_name: gb_est80:BE893687

seq_documentation_block:
LOCUS BE893687 665 bp mRNA EST 20-OCT-2000
DEFINITION 601436488F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3921669 5',
mRNA sequence.
ACCESSION BE893687
VERSION BE893687.1 GI:10355301
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 665)
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
found through the I.M.A.G.E. Consortium information can be
http://image.llnl.gov
Plate: LLAM9754 row: p column: 22
High quality sequence stop: 665.
Location/Qualifiers
1..665
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3921669"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 138 a 165 c 216 g 146 t
ORIGIN

alignment_scores:
Quality: 801.00 Length: 162
Ratio: 4.944 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-486-167A-2 x BE893687 ..

Align seg 1/1 to: BE893687 from: 1 to: 665

1 MetAlaProLLeLysValGlyAspAlaIleProAlaValGluValPheG 17
|||||
.66 ATGGCCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTTCGA 115
|||||
17 uClyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
116 AGGGAGGCCAGG.ACAAGAGTGAACCTGGCAGCTGTTCAAGGCCAAGA 164
|||||

34 ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
165 AGGGTGTCTGTTGGAGTTCCTGGGGCCCTTACCCCTGGATGTTCCAAG 214
|||||
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67
|||||
215 ACACACCTGCCAGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 264
|||||
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
265 AGTCCAGGTGGTGGCCTGCTGAGTGTATGATGCTTTGTGACTGGCG 314
|||||
84 LuTriPGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
315 AGTGGGGCCGAGCCACAGGCGGAAGGCTTGGCTCTCTGGCTGTAT 364
|||||
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
|||||
365 CCCACTGGGGCCTTTGGGAAGGAGACAGACTTATTACTAGATGATCGCT 414
|||||
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
|||||
415 GGTGTCCATCTTTGGGAATCGACGCTCAAGAGGTTCTCCATGGTGGTAC 464
|||||
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
465 AGGATGGCATAGTAGAGGCCCTGAATGTGGAACAGATGGCAGGCGCTC 514
|||||
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
|||||
515 ACCTGCAGCGCTGGCACCACCAATATCATCTCACAGCTC 550
|||||
seq_name: gb_est92:BF792903

seq_documentation_block:
LOCUS BF792903 716 bp mRNA EST 12-JAN-2001
DEFINITION 602253206F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4345488 5',
mRNA sequence.
ACCESSION BF792903
VERSION BF792903.1 GI:12097888
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 716)
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9965 row: 1 column: 01
High quality sequence stop: 623.
Location/Qualifiers
1..716
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4345488"
/clone_lib="NIH_MGC_84"
/tissue_type="adrenal cortex carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
primed. Average insert size 1.229 kb. Library enriched for
full-length clones and constructed by Life Technologies."

BASE COUNT 168 a 174 c 223 g 151 t
ORIGIN

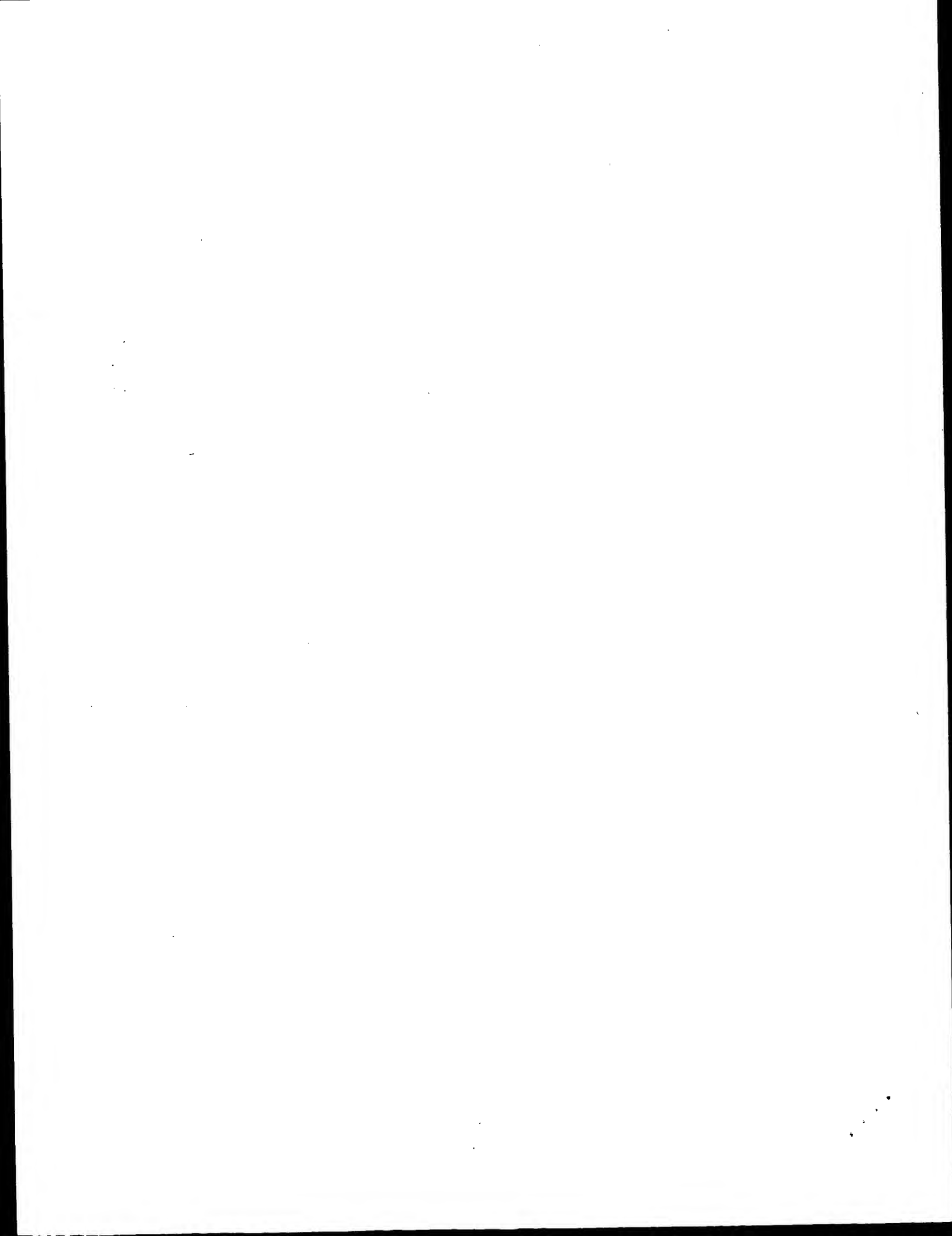
Note: this is a NIH_MGC Library."

alignment_scores:
Quality: 801.00 Length: 162
Ratio: 4.944 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-486-167A-2 x BF792903 ..

Align seg 1/1 to: BF792903 from: 1 to: 716

```
1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
|||||
51 ATGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 100
|||||
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
101 AGGGAGCCAGG.AACAAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 149
|||||
34 ySglyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
150 AGGGTGTGCTGTTGGAGTTCCTGGGGCCTTCACCCCTGGATGTTCAAG 199
|||||
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
|||||
200 ACACCTGCCAGGGTTTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 249
|||||
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
250 AGTCCAGGTGGTGGCTGCTGAGTGTATGATGCTTTGTGACTGGCG 299
|||||
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
300 AGTGGGGCCGAGCCACAAAGCGGAAGGTTCTGGCTCTCTGGCTGAT 349
|||||
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
|||||
350 CCCACTGGGGCCTTTGGGAGGAGACAGACTTATTACTAGATGATCGCT 399
|||||
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
|||||
400 GGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTTCCTCATGGTGTAC 449
|||||
134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
450 AGGATGGCATAGTGAAGGCCCTGAAATGTGTGAACCAAGATGGCACAGGCCTC 499
|||||
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
|||||
500 ACCTGCAGGCTGGCACCAATATCATCTCACAGCTC 535
```



OM of: US-09-486-167A-2 to: N_Geneseq_0601.* out_format : pfs
 Date: Oct 22, 2001 10:44 AM
 About: Results were produced by the GenCore software, version 4.5.
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame2.p2n.model -DEV=xlpl
 -O=cnr2_1/USPTO.spool/US09486167/runat_22102001_064207_25068/app_query.fasta_1.222
 -DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
 -GAPOP=4.500 -GAPEXT=0.050 -GAPOP=10.000 -GAPEXT=0.500
 -GAPOP=6.000 -GAPEXT=7.000 -GAPOP=10.000 -GAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM-ext -MINLEN=0 -MAXLEN=2000000000
 -USER=US09486167@cgn1_1_197 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:

Query: US-09-486-167A-2

Query length: 162

Database: N_Geneseq_0601.*

Database sequences: 730101

Database length: 313950809

Search time (sec): 109.540000

score_list:

Sequence	Strd Orig	zscore	EScore	Len	Documentation
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	830.00	1773.91	7.7e-91	805	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	830.00	1771.67	1.0e-90	993	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	867.00	1427.88	1.5e-71	452	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	573.00	1221.39	4.1e-60	601	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	568.00	1210.53	1.9e-59	604	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	476.00	996.64	1.5e-47	2710	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	423.00	897.98	4.8e-42	553	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	364.00	772.55	4.6e-35	453	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	349.00	739.75	3.1e-33	469	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	315.00	667.34	3.4e-29	423	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	294.00	616.14	2.4e-26	727	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	289.00	614.28	3.0e-26	315	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	289.00	605.19	9.7e-26	737	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	288.00	603.98	1.1e-25	674	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	286.00	603.09	1.3e-25	489	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	286.00	599.52	2.0e-25	683	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	277.00	580.14	2.4e-24	678	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	270.00	563.27	2.1e-23	797	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	259.50	546.15	1.9e-22	473	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	250.50	500.33	6.7e-20	5556	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	250.50	456.00	2.0e-17	34980	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	247.50	514.73	1.1e-20	789	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	241.50	455.61	2.1e-17	58909	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	238.50	494.97	1.3e-19	811	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	236.00	489.58	2.7e-19	812	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	227.00	471.62	2.7e-18	705	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	219.50	371.28	1.0e-12	183012	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	218.50	455.72	2.1e-17	559	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	214.50	373.63	7.7e-13	534720	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	214.50	373.60	7.7e-13	536165	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	197.50	416.29	3.2e-15	320	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	194.00	404.05	1.6e-14	435	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	194.00	402.91	1.8e-14	551	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	194.00	402.85	1.8e-14	554	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	191.50	399.39	2.8e-14	462	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	191.50	395.19	4.1e-14	618	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	173.50	352.30	1.2e-11	991	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	128.50	260.38	1.6e-06	600	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	108.00	178.74	0.0551	19619	
/SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965 +	101.00	201.58	0.0029	564	

/SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AA25092 + 97.50 196.67 0.0055 440
 /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AA25092 + 97.50 190.14 0.0128 810
 /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AA25092 + 97.00 191.40 0.0109 651
 /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT:AA251788 + 96.00 185.56 0.0230 918
 /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AA251788 + 96.00 185.56 0.0230 918

seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AA27965

seq_documentation_block:

ID_AA27965 standard; cDNA: 805 BP.

AC_AA27965;

DT_08-JUN-1999 (first entry)

XX_Human bronchoalveolar polypeptide, B18hum, coding sequence.

DE_B18hum: bronchoalveolar protein; peroxisome-associated polypeptide;
 lung injury; oxidative stress-related disorder; inflammatory disease;
 cardiovascular disease; neurodegenerative disorder; allergic reaction;
 amyotrophic lateral sclerosis; apoptosis; high bone mass syndrome;
 osteopetrosis; osteoporosis-pseudoglioma syndrome; diagnosis;
 Bardet-Biedl syndrome 1; therapy; ss.

OS_Homo sapiens.

PN_WO9909054-A2.

PD_25-FEB-1999.

PF_20-AUG-1998; 98WO-BE00124.

PR_20-AUG-1997; 97BE-0000692.

XX_ (UYLO-) UNIV CATHOLIQUE LOUVAIN.

PA_ (UYMO-) UNIV MONS-HAINAUT.

PI_Bernard A, Falmagne P, Hermans C, Knoop B, Wattiez R;

DR_WPI; 1999-180968/15.

XX_P-PSDB; AAY01079.

XX_New low molecular weight human broncho-alveolar polypeptide -
 useful for diagnosis and/or treatment of lung injuries and diseases,
 and oxidative stress-related diseases and disorders, especially
 inflammatory diseases

PS_Claim 8; Page 33-34; 45pp; English.

XX_This sequence encodes the human bronchoalveolar polypeptide, designated
 B18hum of the invention. B18hum is a low molecular weight human,
 peroxisome-associated broncho-alveolar polypeptide. A diagnostic device
 featuring the polypeptide, polynucleotide and/or inhibitor is useful for
 in vitro detection of lung injuries and diseases or oxidative
 stress-related diseases and disorders, especially inflammatory diseases.
 The device is also useful for monitoring such diseases or disorders in
 patients or fluid samples. The polypeptide, polynucleotide and/or inhibitor
 form pharmaceutical compositions useful in the prevention and/or
 treatment of these diseases or disorders, especially specific
 cardiovascular diseases (e.g. atherosclerosis), neurodegenerative
 disorders (e.g. Alzheimer's disease and Parkinson's disease), amyotrophic
 lateral sclerosis, apoptosis and inflammatory reactions, allergic
 reactions (e.g. asthma, hay fever and eczema), high bone mass syndrome,
 osteopetrosis, osteoporosis-pseudoglioma syndrome, and Bardet-Biedl
 syndrome 1. The polypeptide is also useful as a specific marker of the
 above diseases or disorders in a wide variety of tissues.
 The discovery of the peroxisome-associated polypeptide
 enables the development of diagnosis and treatment of peroxisomal
 disorders.

XX_Sequence 805 BP; 163 A; 200 C; 276 G; 166 T; 0 other;

alignment_scores:

Quality: 830.00 Length: 162
Ratio: 5.123 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-486-167A-2 x AAX27965 ..

Align seg 1/1 to: AAX27965 from: 1 to: 805

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
193 ATGGCCCCAATCAAGTGGAGATGCTATCCAGCAGTGGAGGTGTTGA 242
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
243 AGGGAGCCAGGAGCAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 292
34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
293 AGGGTGTGCTGTTGGAGTTCCTGGGGCCTTCACCCCTGGATGTTCCAAG 342
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
343 ACACACTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 392
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
393 AGTCCAGGTGGTGGCTGCTCAGTGTAAATGATGCTTTGTGACTGGCG 442
84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
443 AGTGGGGCCGAGCCCAAGAGGGGAAGGCAAGGTTCTGGCTCGCTGAT 492
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
493 CCCACTGGGGCTTTGGGAGAGAGACACTTATTACTAGATGATTCGCT 542
117 uValSerLeuPheGlyAsnArgLeuLysArgPheSerMetValValG 134
543 GGTGCTCATCTTTGGGAATCGACGCTCTCAAGAGGTTCTCCATGGTGTAC 592
134 InAspGlyLeuValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
593 AGGATGGCATAGTAGAGGCCCTGAAATGTGGAACACGATGGCACAGGCCTC 642
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
643 ACCTCGAGCTGGCACCCCAATATCATCTCACAGCTC 678

seq_name: /SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT: AAX56411

seq_documentation_block:

ID AAX56411 standard; cDNA; 993 BP.

AC AAX56411;

DT 22-JUL-1999 (first entry)

DE Human vesicle membrane protein-like protein 1 encoding cDNA.

KW Human; vesicle membrane protein-like protein; VMP; Cushing's syndrome;
KW developmental disorder; vesicle-trafficking disorder; cystic fibrosis;
KW immunological disorder; reproductive disorder; neoplastic disorder;
KW anaemia; muscular dystrophy; cataract; Grave's disease; allergy;
KW ulcerative colitis; microbial infection; Addison's disease; cancer; ss.

OS Homo sapiens.

XX 19921994-A2.

XX 06-MAY-1999.

XX 4.

PF 14-OCT-1998; 98WO-US21730.
XX
PR 28-OCT-1997; 97US-0959004.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Corley NC, Hillman JL, Lal P, Shah P, Yue H;
XX
DR WPI; 1999-326702/27.
XX
PT P-PSDB; AAY17388.

Human vesicle membrane-like proteins, useful for diagnosis,
treatment and prevention of e.g. developmental disorders

XX Claim 7; Fig 1; 105pp; English.

XX The present sequence encodes human vesicle membrane protein-like
CC protein 1 (VMP-1). VMP proteins, and their fragments, are used to treat
CC or prevent developmental or vesicle-trafficking disorders, while their
CC antagonists are used to treat or prevent immunological, reproductive or
CC neoplastic disorders. Typical of many such disorders are anaemia;
CC Cushing's syndrome; muscular dystrophy; cataract; cystic fibrosis;
CC Grave's disease; ulcerative colitis; allergies; microbial infections;
CC Addison's disease; cancer of breast, testis and prostate. VMP proteins
CC are also used to raise specific antibodies (used to detect VMP in
CC immunoassays (for diagnosis or monitoring), in competitive drug screens
CC and to purify VMP from natural sources) and to screen for specific
CC antagonists (potential therapeutic agents). VMP polynucleotides or their
CC fragments, are used in hybridization assays to detect VMP in biological
CC samples (e.g. for diagnosis, including detection of mutations and
CC polymorphisms), optionally after amplification; to express recombinant
CC VMP, including in vivo for gene therapy, and to map the corresponding
XX genomic sequence.

SQ Sequence 993 BP; 176 A; 260 C; 358 G; 197 T; 2 other;

alignment_scores:

Quality: 830.00 Length: 162
Ratio: 5.123 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-486-167A-2 x AAX56411 ..

Align seg 1/1 to: AAX56411 from: 1 to: 993

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
385 ATGGCCCCAATCAAGTGGAGATGCTATCCAGCAGTGGAGGTGTTGA 434
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
435 AGGGAGCCAGGAGCAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 484
34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
485 AGGGTGTGCTGTTGGAGTTCCTGGGGCCTTCACCCCTGGATGTTCCAAG 534
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
535 ACACACTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 584
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
585 AGTCCAGGTGGTGGCTGCTGAGTGTAAATGATGCTTTGACTGGCGG 634
84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
635 AGTGGGGCCGAGCCCAAGGGGAAGGCAAGGTTCTGGCTCGCTGAT 684
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
|||||

685 CCCATGGGGCCCTTTGGGAAGGAGACAGACTATTACTAGATGATTCCT 734
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
 |||||
 735 GGTGTCCTCTTTGGGAATCGACGCTCTCAAGAGGTTCTCCATGGTGTAC 784
 134 luAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
 |||||
 785 AGGATGGCATAGTAGAAGGCCCTGTAATGTGAACACAGATGGCACAGCCCTC 834
 151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
 |||||
 835 ACTCGAGCTGGCAGCCCAATATATCATCTACAGAGCTC 870

seq_name: /SIDS1/gcgdata/geneseq/NA2001.DAT:AAF92316

seq_documentation_block:

ID AAF92316 standard; cDNA; 452 BP.

AC AAF92316;

DT 15-MAY-2001 (first entry)

DE Bovine mammary tissue derived cDNA #29.

KW Bovine; mammary gland; cancer; tumour; angiogenesis; ss.

OS Bos taurus.

PN WO200114553-A1.

PD 01-MAR-2001.

PF 23-AUG-2000; 2000WO-NZ00166.

PR 23-AUG-1999; 99US-0150330.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PA (NZPA-) NEW ZEALAND PASTORAL AGRIC RES INST LTD.

PI Havukkala IJ, Gleen M, Grigor MR, Molenaar AJ;

DR WPI; 2001-226619/23.

PT New polypeptides and polynucleotides encoding the polypeptides, which are expressed in bovine mammary gland tissue, useful for stimulating mammary gland growth or function, or inducing differentiation of milk producing cells -

PS Claim 1; Page 49-50; 97pp; English.

CC The present invention relates to proteins derived from bovine mammary gland cells. The invention is useful for stimulating bovine mammary gland cell growth and function, inhibiting the growth of various mammary gland cancer cells, inhibiting angiogenesis and vascularization of tumours, or modulating the growth of blood vessels in a mammal.

SQ Sequence 452 BP; 100 A; 103 C; 145 G; 104 T; 0 other;

alignment_scores:

Quality: 667.00 Length: 146

Ratio: 4.697 Gaps: 0

Percent Similarity: 97.260 Percent Identity: 86.986

alignment_block:

US-09-486-167A-2 x AAF92316 ..

Align seg 1/1 to: AAF92316 from: 1 to: 452

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
 |||||

14 ATGCCCCCGGATTAAAGTTGGAGATGCCATTCCATCGGTGGAGGATTATGA 63
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
 |||||
 64 AARGAGCCCGGCAACAAGGTGAACCTGGCAGAGCTGTTCAAGGCAAGA 113
 34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
 |||||
 114 AGGAGTGCCTGTTTGGCTCCCTGGGCGCTTACCCCTGGTGGTTCAG 163
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
 |||||
 164 ACCACCTGCCAGGTTTCGTGGAGCAGGCTGATGCTCTGAAGGCCAAGG 213
 67 yValGlnValAlaAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 |||||
 214 GATCCAGTGGTGGCATGCTGACCGTTAATGATGCTCTTGTAACTGAAG 263
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
 |||||
 264 AGTGGCAGCGCCCAAGCAGAGGCAAGGTTCCGGCTCCTGGCAGAC 313
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117
 |||||
 314 CCCAGTGGGACTTTGGGAAGGACAGATTGTTACTTGATGATTCAC 363
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
 |||||
 364 GCTCTTCTCTTTGGGAATCACCAGCTGAAGAGGTTCTCCATGGTGATAG 413
 134 luAspGlyIleValLysAlaLeuAsnValGluProAsp 146
 |||||
 414 AGGATGCATCGTCAATCCCTGAACCTGGAGCCAGAT 451

seq_name: /SIDS1/gcgdata/geneseq/NA1999.DAT:AAx27966

seq_documentation_block:

ID AAX27966 standard; cDNA; 601 BP.

AC AAX27966;

DT 08-JUN-1999 (first entry)

DE Human bronchoalveolar polypeptide, B18hum, coding sequence.

KW B18hum; bronchoalveolar protein; peroxisome-associated polypeptide; lung injury; oxidative stress-related disorder; inflammatory disease; cardiovascular disease; neurodegenerative disorder; allergic reaction; amyotrophic lateral sclerosis; apoptosis; high bone mass syndrome; osteopetrosis; osteoporosis-pseudoglioma syndrome; diagnosis; Bardet-Biedl syndrome 1; therapy; ss.

OS Homo sapiens.

XX WO9909054-A2.

PN 25-FEB-1999.

PF 20-AUG-1998; 98WO-BE00124.

PR 20-AUG-1997; 97BE-0000692.

PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.

PA (UYMO-) UNIV MONS-HAINAUT.

PI Bernard A, Falmagne P, Hermans C, Knoops B, Wattiez R;

DR WPI; 1999-180968/15.

XX New low molecular weight human broncho-alveolar polypeptide - useful for diagnosis and/or treatment of lung injuries and diseases, PT and oxidative stress-related diseases and disorders, especially PT inflammatory diseases

XX PS Disclosure; Page 40; 45pp; English.
 XX CC This sequence encodes the human bronchoalveolar polypeptide, designated
 CC B18hum of the invention. B18hum is a low molecular weight human,
 CC peroxisome-associated broncho-alveolar polypeptide. A diagnostic device
 CC featuring the polypeptide, polynucleotide and/or inhibitor is useful for
 CC in vitro detection of lung injuries and diseases or oxidative
 CC stress-related diseases and disorders, especially inflammatory diseases.
 CC The device is also useful for monitoring such diseases or disorders in
 CC patients or fluid samples. The polypeptide, polynucleotide and inhibitor
 CC form pharmaceutical compositions useful in the prevention and/or
 CC treatment of these diseases or disorders, especially specific
 CC cardio-vascular diseases (e.g. atherosclerosis), neurodegenerative
 CC disorders (e.g. Alzheimer's disease and Parkinson's disease), amyotrophic
 CC lateral sclerosis, apoptosis and inflammatory reactions, allergic
 CC osteopetrosis, osteoporosis-pseudoglioma syndrome, high bone mass syndrome,
 CC syndrome 1. The polypeptide is also useful as a specific marker of the
 CC above diseases or disorders in a wide variety of tissues.
 CC The discovery of the peroxisome-associated polypeptide
 CC enables the development of diagnosis and treatment of peroxisomal
 CC disorders.
 XX SQ Sequence 601 BP; 115 A; 162 C; 200 G; 124 T; 0 other;

alignment_scores:
 Quality: 573.00 Length: 113
 Ratio: 5.071 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-486-167A-2 x AAX27966 ..

Align seg 1/1 to: AAX27966 from: 1 to: 601

50 LysThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLys 66
 |||||
 173 AAGACACACCTCCAGGGTTTGTGGAGCAGGCTGAGGCTCTCAAGGCCAA 222
 66 sclyValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrG 83
 |||||
 223 GGGAGTCCAGGTTGGTGGCTGTCTGAGTGTAAATGATGCTTTGTGACTG 272
 83 LysGluTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuAla 99
 |||||
 273 GCGAGTGGGCGCGAGCCACAGCGCGAAGCAAGGTCGGCTCTGGCT 322
 100 AspProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspAspSe 116
 |||||
 323 GATCCCACTGGGCGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTC 372
 116 rleuValSerilePheGlyAsnArgArgLeuLysArgPheSerMetValV 133
 |||||
 373 GCTGGTGTCATCTTTGGAAATTCGACGCTCAAGAGGTTCTCCATGGTGG 422
 133 alGlnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGly 149
 |||||
 423 TACAGATGGCATATGAAGGCCCTGAATGTGGACACAGATGGCACAGGC 472
 150 LeuThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
 |||||
 473 CTCACCTGCAGCGCTGGCACCACCAATATCATCTCACAGCTC 511

seq_name: /SID1/gcdata/geneseq/geneseq/NAL99.DAT: AAX27968

seq_documentation_block:

ID AAX27968 standard; cDNA; 604 BP.

XX AAX27968;

XX 08-JUN-1999 (first entry)

DT

XX DE Human bronchoalveolar polypeptide, B18hum, coding sequence.
 XX KW B18hum; bronchoalveolar protein; peroxisome-associated polypeptide;
 KW lung injury; oxidative stress-related disorder; inflammatory disease;
 KW cardiovascular disease; neurodegenerative disorder; allergic reaction;
 KW amyotrophic lateral sclerosis; apoptosis; high bone mass syndrome;
 KW osteopetrosis; osteoporosis-pseudoglioma syndrome; diagnosis;
 KW Bardet-Biedl syndrome 1; therapy; ss.
 XX OS Homo sapiens.
 XX PN WO9909054-A2.
 XX PD 25-FEB-1999.
 XX PF 20-AUG-1998; 98WO-BE00124.
 XX PR 20-AUG-1997; 97BE-0000692.
 XX PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.
 XX PA (UYMO-) UNIV MONS-HAINAUT.
 XX PI Bernard A, Falmagne P, Hermans C, Knoop B, Wattiez R;
 XX WPI; 1999-180968/15.
 XX PT New low molecular weight human broncho-alveolar polypeptide -
 PT useful for diagnosis and/or treatment of lung injuries and diseases,
 PT and oxidative stress-related diseases and disorders, especially
 PT inflammatory diseases
 XX PS Disclosure; Page 40-41; 45pp; English.
 XX CC This sequence encodes the human bronchoalveolar polypeptide, designated
 CC B18hum of the invention. B18hum is a low molecular weight human,
 CC peroxisome-associated broncho-alveolar polypeptide. A diagnostic device
 CC featuring the polypeptide, polynucleotide and/or inhibitor is useful for
 CC in vitro detection of lung injuries and diseases or oxidative
 CC stress-related diseases and disorders, especially inflammatory diseases.
 CC The device is also useful for monitoring such diseases or disorders in
 CC patients or fluid samples. The polypeptide, polynucleotide and inhibitor
 CC form pharmaceutical compositions useful in the prevention and/or
 CC treatment of these diseases or disorders, especially specific
 CC cardio-vascular diseases (e.g. atherosclerosis), neurodegenerative
 CC disorders (e.g. Alzheimer's disease and Parkinson's disease), amyotrophic
 CC lateral sclerosis, apoptosis and inflammatory reactions, allergic
 CC reactions (e.g. asthma, hay fever and eczema), high bone mass syndrome,
 CC osteopetrosis, osteoporosis-pseudoglioma syndrome, and Bardet-Biedl
 CC syndrome 1. The polypeptide is also useful as a specific marker of the
 CC above diseases or disorders in a wide variety of tissues.
 CC The discovery of the peroxisome-associated polypeptide
 CC enables the development of diagnosis and treatment of peroxisomal
 CC disorders.
 XX SQ Sequence 604 BP; 117 A; 160 C; 200 G; 127 T; 0 other;

alignment_scores:
 Quality: 568.00 Length: 162
 Ratio: 4.814 Gaps: 1
 Percent Similarity: 72.840 Percent Identity: 72.840
 alignment_block:
 US-09-486-167A-2 x AAX27968 ..
 Align seg 1/1 to: AAX27968 from: 1 to: 604
 1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
 |||||
 161 ATGGCCCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 210
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34


```
51 ..... 51
1252 GGCAGAGAAATGCTTGAACCCAGAGGGTAGTTGCAGTCAGTGGAGAT 1301
51 ..... 51
1302 CGTGCCTCTGCAGTCCAGCCTGGGTGAAGAGCGAGACTCGGTCTCAAAA 1351
51 ..... 51
1352 TGAATAAATAAAGAAAGAAAGTAGAGACTGCAAAAGGAACAGTACC 1401
51 ..... 51
1402 GGGATGTGTGAGAAACATACATACATTAATCCACACCCCTGTTGG 1451
51 ..... 51
1452 TCCTGCTAAATGACAGGCACCTGTGGAAGGTGCTTGGGACTCAGATAATA 1501
51 ..... 51
1502 AGACAAGATCTGCCATGGAAAGTTACGCTCTGGACCATTAAGCATTAG 1551
51 ..... 51
1552 GTTTCATTTCTGAGCTTCTAGTGCCCAAGGCAAAAGGAATAGATGGT 1601
51 ..... 51
1602 TTAGACAGCTCTCATTTGCTGATCAAGGTGTTGAGGCAGACACTGAGG 1651
51 ..... 51
1652 AGGGCTGGAGATAAAGGTGGGTGGGGTGCAGATGCAGTTATCCCTTT 1701
52 .....His..LeuProGlyPheValGlu 58
1702 GCCGACCCTTTGTCCTCCCTCAGACACACCTGCCAGGCTTTGTGGAG 1751
59 GlnAlaGluAlaLeuLysAlaLysGlyValGlnValValAlaCysLeuSe 75
1752 CAGGCTGAGGCTCTGAAGCCCAAGGGAGTCCAGGTGGCTGTCTGTGAG 1801
75 rValAsnAspAlaPheValThrGlyGluTrpGlyArgAlaHisLysAlaG 92
1802 TGTAAATGATGCTTTGTGACTGGCGAGTGGGGCCGAGCCACAAAGCGG 1851
92 luGlyLys..... 94
1852 AAGGCAAGGTGAGGTGAGGGGCTGCAGGAGTGCAGGACCAGGTAGGATA 1901
95 .....ValArgLeuLeuAlaAspPr 101
1902 TTTCTTTTGTGACCTCTACTTTCTGAGGTTTCGGCTCCTGGCTGATCC 1951
101 oThrGlyAlaPhe.GlyLys..... 107
1952 CACTGGGCGCTTTGGGAAGGTGAGTGTCCCTGTGACCCGACAGGGACAT 2001
107 ..... 107
2002 GCGGTGCGGGAGCAGTGGGGCCCTTGGCCCTCTCAAGGATTTCTGCAC 2051
108 .....GluThrAspLeuLeuLeuAspSerL 117
2052 ACTTTTCTCTGTCTCTCTTAGGACACAGACTATTACTAGATGATCGC 2101
117 euValSerIlePheGlyAsnArgArgLeu..Lys..... 127
2102 TGGTGTCCATCTTTGGGAATCGACCTCTCAAGAGGTAAAGTGGAGAGTC 2151
```

```
127 ..... 127
2152 CTCTGTGGAGAAAGTCTCTCTGTGGGAGAGAGTCCTCTGTGGGAGAGATC 2201
127 ..... 127
2202 CTCTGTGGAGAGGTCCTCTCTGTGGAAAGAGTCGTCGTGTGGGGGAGATGTG 2251
127 ..... 127
2252 TGGGAGAGAGTCCTCTGTGTGGGAGAGTCTTCTGTAGGGGAGAGTCCCTCTG 2301
127 ..... 127
2302 GGGAGAGAGTCTCTGTGTGGGAGAGTCTCTGTGTGGGAGAGTCCCTCTGT 2351
127 ..... 127
2352 GTGGAGAGAGTCCTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2401
127 ..... 127
2402 GTGGGGGAGTCCTCTCTGTGGAGTCTCTTGTGGCCCTGGCTGTTCACATGC 2451
127 ..... 127
2452 CTGTCTCCATGCCAGCTCCCAAGCCAGCTGATGCAGCTGGCTGGGCC 2501
128 .....ArgPheSerMetValValGlnAspGlyIleValLysAl 140
2502 CCTCTTCCGGCAGGTTCTCCATGGTGTACAGGATGGCATAGTAGTAAGGC 2551
140 aLeuAsnValGluProAspGlyThrGlyLeuThrCysSerLeuAlaProA 157
2552 CCTGAATGTGAACACAGATGCACAGGCTCACCTGCAGCCTGGCACCCA 2601
157 snIleIleSerGlnLeu 162
2602 ATATCATCTCACAGCTC 2618
seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AAC95054
seq_documentation_block:
ID AAC95054 standard; cDNA; 553 BP.
XX
AC AAC95054;
XX
DT 19-FEB-2001 (first entry)
XX
DE Cat flea hindgut and Malpighian tubule (HMT) cDNA, SEQ ID NO:1549.
XX
KW Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;
KW flea infestation; vaccine; antiparasitic; therapeutic target;
KW diagnosis; detection; ss.
XX
OS Ctenocephalides felis.
XX
PN WO200061621-A2.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000WO-US09437.
XX
PR 09-APR-1999; 99US-0128704.
XX
PA (HESK-) HESKA CORP.
XX
PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;
XX
DR WPI; 2000-656323/63.
XX
PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic
acids useful for the prevention, diagnosis and treatment of flea
```


PT infestations -

XX Claim 26; Page 756; 964pp; English.

XX The invention relates to novel cat flea (*Ctenocephalides felis*) nucleic acids which are expressed in hindgut and Malpighian tubule (HMT) tissue or head and nerve cord (HNC) tissue. The invention also relates to the encoded proteins. The invention additionally encompasses expression constructs, recombinant viruses and recombinant cells comprising the nucleic acids of the invention, recombinant production of the proteins, antibodies against the proteins, a method of identifying inhibitors of the proteins, and compositions comprising the inhibitors for administration to an animal. The nucleic acids, and the proteins they encode may be used in the prevention, treatment and diagnosis of diseases associated with flea infestations. For example, the nucleic acids may be used to produce an HMT or HNC protein according to standard recombinant DNA methodology by inserting the nucleic acids into a host cell and culturing the cell to express the protein. The HMT and HNC nucleic acids may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect and quantitate the presence of cat flea or other homologous nucleic acid sequences in samples. They may also be used to study the expression and function of the proteins and their role in metabolism. The HMT and HNC proteins may be used as antigens in the production of specific antibodies, and in assays to identify modulators (agonists and antagonists) of HMT and/or HNC protein expression and activity. The anti-HMT/HNC protein antibodies and antagonists may also be used to down-regulate protein expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of flea polypeptides in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The present sequence represents a cat flea HMT cDNA of the invention.

XX Sequence 553 BP; 180 A; 86 C; 123 G; 164 T; 0 other;

alignment_scores:

Quality: 423.00 Length: 164
Ratio: 3.439 Gaps: 3
Percent Similarity: 75.000 Percent Identity: 55.488

alignment_block:

US-09-486-167A-2 x AAC95054 ..

Align seg 1/1 to: AAC95054 from: 1 to: 553

1 MetAlaProLysValGlyAspAlaIleProAlaValGluValPheG1 17
24 ATGGCTCAGATTAAAGTAGGCGATAAAATTCGGTCGGTAGACTATTGGA 73
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
74 AGACATCCAGCTAACAAAGTGAATATTGCAAAATCTCGCTGCAGGAAAA 123
34 ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
124 AAGTGTCTTATTTCGCTGTTCTTGGCCCTTCACTCCAGGATGTTCTAG 173
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
174 ACTCATGTACTGCTGCTATGTTGCAAGGCTGAAGAATTAAGAGAGTGG 223
67 yValGlnValAlaLysLeuSerValAsnAspAlaPheValThrClyG 84
224 AATTGCTGAATCTTTGGCTCTCTGTATGATFCCTTTGTTATGATGATG 273
84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
274 CTTGGGGTAAGGATCAGCAGAGCAATGAAAGGTAAGATGCTTGCAGAT 323
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerL 117
324 CCAAGTGGTACATTCAAAAAGAACTTGGATGGAGTTG..... 363
117 euValSerIlePheGlyAsnArgArgLeuLysArgPhe..... 129

364ATTCGCCCTTAGGAGTTTACGTTCTAAGAG 396
130 ...SerMetValValGlnAspGlyLysValLysAlaLeuAsnValGluPr 145
397 ATCTATGTTGATGACAAATGTCGTATCAGAAATGTAATGTAGAACC 446
145 oAspGlyThrGlyLeuThrCysSerLeuAlaProAsnIle 158
447 AGATGGTCTTGGACTTTCGTTCTTACGTCGACAAACTT 485

seq_name: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT:AAC04049

seq_documentation_block:

ID AAC04049 standard; cDNA; 453 BP.

XX AAC04049;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 4047.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GBST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG04043.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 4047; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX mRNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors.

XX Sequence 453 BP; 82 A; 108 C; 180 G; 82 T; 1 other;

alignment_scores:

Quality: 364.00 Length: 72
Ratio: 5.127 Gaps: 0
Percent Similarity: 98.611 Percent Identity: 98.611

alignment_block:

US-09-486-167A-2 x AAC04049 ..

Align seg 1/1 to: AAC04049 from: 1 to: 453

1 MetalapProIIeLysValGlyAspAlaIleProAlaValGluValPheG1 17
226 ATGCCCCCAATCAAGGTGGAGATGCCATCCAGCAGTGGAGGTGTTGA 275
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
276 AGGGAGCCAGGACACAGGTGACCTGGCAGACCTGTTCAAGGGCAGA 325
34 ySgLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
336 AGGGTGTCTGTTGGAGTCTCTGGGGCTTCAACCTGGATGTTCCAA 375
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysG1 67
376 ACACACCTGCCAGGGTTTGTGGAGCAGCTGAGGCTCTGAAGGCCAAGG 425
67 yValGlnValValAla 72
426 AGTCCAGGTGGTGGCT 441

seq_name: /SISL/gcdata/geneseq/geneseq/NA1999.DAT:AAx27967

seq_documentation_block:

ID_AAX27967 standard; cDNA; 469 BP.

XX AAX27967;

08-JUN-1999 (first entry)

Human bronchoalveolar polypeptide, B18hum, coding sequence.

B18hum; bronchoalveolar protein; peroxisome-associated polypeptide;
lung injury; oxidative stress-related disorder; inflammatory disease;
cardiovascular disease; neurodegenerative disorder; allergic reaction;
amyotrophic lateral sclerosis; apoptosis; high bone mass syndrome;
osteopetrosis; osteoporosis-pseudoglioma syndrome; diagnosis;
Bardet-Biedl syndrome 1; therapy; ss.

XX Homo sapiens.

PN WO9909054-A2.

XX 25-FEB-1999.

PF 20-AUG-1998; 98WO-BE00124.

XX 20-AUG-1997; 97BE-0000692.

XX (UVO-) UNIV CATHOLIQUE LOUVAIN.

PA (UYMO-) UNIV MONS-HAINAUT.

XX Bernard A, Falmagne P, Hermans C, Knoop B, Wattiez R;

XX WPI; 1999-180968/15.

XX New low molecular weight human broncho-alveolar polypeptide
PT useful for diagnosis and/or treatment of lung injuries and diseases,
PT and oxidative stress-related diseases and disorders, especially
PT inflammatory diseases

XX Disclosure; Page 39; 45pp; English.

XX This sequence encodes the human bronchoalveolar polypeptide, designated
CC B18hum of the invention. B18hum is a low molecular weight human,
CC peroxisome-associated broncho-alveolar polypeptide. A diagnostic device
CC featuring the polypeptide, polynucleotide and/or inhibitor is useful for
CC in vitro detection of lung injuries and diseases or oxidative
CC stress-related diseases and disorders, especially inflammatory diseases.
CC The device is also useful for monitoring such diseases or disorders in
CC patients or fluid samples. The polypeptide, polynucleotide and inhibitor

CC form pharmaceutical compositions useful in the prevention and/or
CC treatment of these diseases or disorders, especially specific
CC cardio-vascular diseases (e.g. atherosclerosis), neurodegenerative
CC disorders (e.g. Alzheimer's disease and Parkinson's disease), amyotrophic
CC lateral sclerosis, apoptosis and inflammatory reactions, allergic
CC reactions (e.g. asthma, hay fever and eczema), high bone mass syndrome,
CC osteopetrosis, osteoporosis-pseudoglioma syndrome, and Bardet-Biedl
CC syndrome 1. The polypeptide is also useful as a specific marker of the
CC above diseases or disorders in a wide variety of tissues.
CC The discovery of the peroxisome-associated polypeptide
CC enables the development of diagnosis and treatment of peroxisomal
CC disorders.

XX Sequence 469 BP; 88 A; 132 C; 150 G; 99 T; 0 other;

alignment_scores:

Quality: 349.00 Length: 90

Ratio: 4.653 Gaps: 1

Percent Similarity: 83.333 Percent Identity: 82.222

alignment_block:

US-09-486-167A-2 x AAX27967

Align seg 1/1 to: AAX27967 from: 1 to: 469

83 GlyGluTrp.....GlyArgAlaHi 89

|||||||::|

110 GGAGAGTGGGCTCTGGCGGGTCCGCGAGTTTCACGAGCGCGTGCAGC 159

89 sLysAlaGluGlyLysValArgLeuAlaAspProThrGlyAlaPheG 106

|||||||

160 CATGCCCCCAATCAAGGTTCGGCTCTCGCTGATCCCATCGGGCTTTG 209

106 lYLySGluThrAspLeuLeuAspSerLeuValSerIlePheGly 122

|||||||

210 GGAAGGACACAGACTATTACTAGATGATTCGTGTGTCATCTTTGG 259

123 AsnArgArgLeuLysArgPheSerMetValValGlnAspGlyIleVally 139

|||||||

260 AATCGACGCTCAAGAGGTTCTCCATGCTGTGTACAGGATGGCATAGTAA 309

139 sAlaLeuAsnValGluProAspGlyThrGlyLeuThrCysSerLeuAlap 156

|||||||

310 GGCCCTGAATGTGGAACAGATGGCACAGCGCTCACCTGCAGCGCTGGCAC 359

156 roAsnIleIleSerGlnLeu 162

|||||||

360 CCATATATCATCTCACAGCTC 379

seq_name: /SISL/gcdata/geneseq/geneseq/NA2000.DAT:AAx02224

seq_documentation_block:

ID_AAC02224 standard; cDNA; 423 BP.

XX AAC02224;

XX 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 2222.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

```

XX (GSET) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX P-PSDB; AAG02218.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 222; 71pp + CD-ROM; English.
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX mRNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors.
XX Sequence 423 BP; 77 A; 109 C; 156 G; 81 T; 0 other;

alignment_scores:
    Quality: 315.00      Length: 83
    Ratio: 4.632        Gaps: 1
    Percent Similarity: 81.928      Percent Identity: 80.723

alignment_block:
US-09-486-167A-2 x AAC02224 ..
Align seg 1/1 to: AAC02224 from: 1 to: 423

83 GlyGluTrp.....GlyArgAlaHi 89
|||||
175 GGAGAGTGGCGTGGCGGGTCCGAGTTTCACGACAGCCGTCGAGC 224

89 slySalaGluGlyLysValArgLeuLeuAlaAspProThrGlyAlaPheG 106
|||
225 CATGCCCCCAATCAAGTTCGCTGCTGCTGATCCACCTGGGCGCTTG 274

106 lYlYsGluThrAspLeuLeuAspSerLeuValSerIlePheGly 122
|||||
275 GGAGGAGACAGACTTATTAGATGATTCGCTGGTGCATCTTTGGG 324

123 AsnArgArgLeuLysArgPheSerMetValValGlnAspGlyIleVally 139
|||||
325 AATCGACGCTCTCAAGAGTTCTCATGTTGTTACAGGATGGCATGTGAA 374

139 sAlaLeuAsnValGluProAspGlyThrGlyLeuThrCysSerLeuAla 155
|||||
375 GGCCCTGAATGTGAACAGATGCACAGGCGCTCACCTGCAGCGCTGCA 423

seq_name: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT:AAC33030
seq_documentation_block:
ID AAC33030 standard; DNA; 727 BP.
XX AAC33030;
XX AAC33030;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 1530.
XX

```

```

KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 07-MAY-1999; 99US-0132486.
XX 11-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0134941.
XX 21-MAY-1999; 99US-0135124.
XX 24-MAY-1999; 99US-0135353.
XX 25-MAY-1999; 99US-0135629.
XX 27-MAY-1999; 99US-0136021.
XX 28-MAY-1999; 99US-0136392.
XX 01-JUN-1999; 99US-0136782.
XX 03-JUN-1999; 99US-0137222.
XX 04-JUN-1999; 99US-0137528.
XX 07-JUN-1999; 99US-0137502.
XX 08-JUN-1999; 99US-0137724.
XX 10-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 14-JUN-1999; 99US-0138847.
XX 16-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 17-JUN-1999; 99US-0139453.
XX 18-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 21-JUN-1999; 99US-0139763.
XX 22-JUN-1999; 99US-0139817.
XX 23-JUN-1999; 99US-0139859.
XX 23-JUN-1999; 99US-0140353.

```

```
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140895.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143824.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150560.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
Quality: 294.00      Length: 169
Ratio: 2.534         Gaps: 6
Percent Similarity: 68.639 Percent Identity: 41.420

alignment_block:
US-09-486-167A-2 x AAC33030 ..

Align seg 1/1 to: AAC33030 from: 1 to: 727

1 MetAlaProIleLysValGlyAspAlaIlePro.....AlaValGluVa 15
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
54 ATGGCTCCAATTGCTGCGCGATGTCGTCCAGACGGAACCACTTCCTT 103

15 lPhe...GluGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheL 31
|||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
104 CTTTCGATGAGATGATCAACTTCAGACTGTCCTCCGTTCACTCTCTGCGCG 153

31 ysgLyLysLysGlyValLeuPheGlyValProGlyAlaPheThrProGly 47
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
154 CCGTAAAAAGTCATCCCTCTTTGGTGTCTCTGCTGCTTCACTCCACC 203

48 CysSerLysThrHisLeuProGlyPheValGluGluAlaGluAlaLeuLy 64
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
204 TGCAGCATGAAGCATGTTCTCTGTTTCATTGAGAAAGCAGAGAGCTGAA 253
```

```

64 salalysGlyValGlnValAlaCysLeuSerValAsnAspAlaPheV 81
111:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
254 GTCAAGGCTGTGATGATCATATTGCTTTAGTGTGAACGATCCTTTT 303
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
81 alThrGlyGluTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeu 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 TGATGAAGCATGGGGAAGACATACCCAGAGACAAGCATGTGAAGTTT 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
98 LeuAlaAspProThrGlyAlaPhe.....GlyLysGluThrAs 110
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
354 GTAGTGATGGCTGTGGAGATACACACACACCTTCTTGACTTGAGCTTGA 403
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
110 pleuLeuLeuAspSerLeuValSerIlePheGlyAsnArgArgLeuL 127
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
404 C.....CTTAAGGACAAGGCTGTGGTATT.....AGGTCAA 435
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
127 ysArgPheSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
436 GGAGATTCGCTTGTGCTTGATTAACCTTAAGGTGACTGTAGCCCAATCTC 485
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
144 GluProAspGlyThrGlyLeuThrCysSerLeuAlaProAsnIleLeSe 160
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
486 GAA...TCTGGTGGCGGGTTCACAGATTCCAGCGCGGATGATATCCTCGA 532
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
160 rGlnLeu 162
|||||
533 GGCTCTC 539

```

seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT:AA211145

seq_documentation_block:

ID AAT21145 standard; cDNA to mRNA; 315 BP.

AC AAT21145;

XX 01-AUG-1996 (first entry)

DT Human gene signature HUMGS02452.

DE Gene signature; messenger RNA; mRNA; relative abundance; frequency;

XX human; cloning; mapping; non-biased library; diagnosis; detection;

KW cell typing; abnormal cell function; ss.

XX Homo sapiens.

XX WO9514772-A1.

XX 01-JUN-1995.

XX 11-NOV-1994; 94WO-JP01916.

XX 12-NOV-1993; 93JP-0355504.

XX {MATS/} MATSUBARA K.

PA {OKUB/} OKUBO K.

XX Matsubara K, Okubo K;

PI MPI; 1995-206931/27.

XX Identifying gene signatures in 3'-directed human cDNA library - e.g.

PT for diagnosis of abnormal cell function, by preparing cDNA that

PT reflects relative abundance of corresp. mRNA in specific human

PT tissues

XX Claim 1; Page 813; 2245pp; Japanese.

XX A single-stranded DNA (or its complementary strand or the corresp.

CC double-stranded DNA) which comprises one of the 7837 "GS" sequences

CC given in AAT19001-T26837 and which is able to hybridise to part of

CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)

CC sequences were obtained from 3'-directed cDNA libraries prepared

CC from various human tissues; synthesis of cDNA was initiated from the

CC

CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.

XX Sequence 315 BP; 69 A; 82 C; 81 G; 75 T; 8 other;

alignment_scores:

Quality: 289.00 Length: 62
 Ratio: 4.898 Gaps: 0

Percent Similarity: 95.161 Percent Identity: 93.548

alignment_block:

US-09-486-167A-2 x AAT21145 ..

Align seg 1/1 to: AAT21145 from: 1 to: 315

100 AspProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSe 116
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

1 GATCCCACTGGGCGCTTTGGGAAGNGACAGACTTATNACTAGATGATTC 50
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

116 rLeuValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValV 133
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

51 GCTGGTGTCCATCTTTGGGAATCGACGTCTCAAAAGGTTATCCATGGTGG 100
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

133 alGlnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGly 149
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

101 TACAGGATGGCATAGTAGAGGCCCTGAATGTGAACACAGATGGCACAGGC 150
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

150 LeuThrCysSerLeuAlaProAsnIleLeSerGln 161
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

151 CTCACCTGCAGCGCTGGCACCAATATCATCTCACAG 186
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AAC34962

seq_documentation_block:

ID AAC34962 standard; DNA; 737 BP.

XX AAC34962;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 8523.

XX Hybridisation assay; genetic mapping; gene expression control;

XX protein identification; signal transduction pathway;

XX metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.


```
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
  Quality: 289.00      Length: 169
  Ratio: 2.535        Gaps: 6
  Percent Similarity: 67.456  Percent Identity: 40.828

alignment_block:
US-09-486-167A-2 x AAC34962 ..

Align seg 1/1 to: AAC34962 from: 1 to: 737

1 MetAlaProLysLysValGlyAspAlaIlePro.....AlaValGluVa 15
|||||
62 ATGCTCCAAATGCTGTCGGGATGTCGTACACGACGCACTATCTCCTT 111
|||||

15 lPhe...GluGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheL 31
|||||
112 CTTCGATGAGAATCACTTCACTTCACTCGCTCGCTCGCTCGCTCGCG 161
|||||

31 ysGlyLysLysGlyValLeuPheGlyValProGlyAlaPheThrProGly 47
|||||
162 CCGTAAAGAAAGTCTCTCTTGGTCTCTGCTGCTTCACTCCCACT 211
|||||

48 CysSerLysThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLy 64
|||||
212 TGCAGCATGAAGCATGTTCTCGGTTTCATTGAGAACACAGAGAGCTGAA 261
|||||

64 sAlaLysGlyValGlnValValAlaCysLeuSerValAsnAspAlaPheV 81
|||||
262 GTCAAAGGGTGTGATGAGATCAATTTCTTTAGCGTGAACAATCCGTTTG 311
|||||

81 alThrGlyGluTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeu 97
|||||
312 TGATGAAGCATGGGGAACATACACACACACACCTTCTCGGACTTGAGTTG 361
|||||

98 LeuAlaAspProThrGlyAlaPhe.....GlyLysGluThrAs 110
|||||
362 GTAGCTGATGGTCTGGAGAATACACACACCTTCTCGGACTTGAGTTGA 411
|||||

110 pLeuLeuLeuAspSerLeuValSerIlePheGlyAsnArgArgLeuL 127
|||||
412 C.....CTTAAGACAAGGGTCTTGGTGT.....AGGTCAA 443
|||||

127 ysArgPheSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
|||||
444 GGAGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 493
|||||

144 GluProAspGlyThrGlyLeuThrCysSerLeuAlaProAsnIleLeu 160
|||||
494 GAATCTGGTGGCGAA...TTCACAGTTTCCGACGCGCGATGATCTCTCAA 540
|||||
```

```
160 rGlnLeu 162
: |||
541 GGCTCTC 547

seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:AAC45603

seq_documentation_block:
ID AAC45603 standard; DNA; 674 BP.
XX
AC AAC45603;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 47093.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127482.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
```

```
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:      Quality: 288.00      Length: 169
                        Ratio: 2.483          Gaps: 6
Percent Similarity: 68.639      Percent Identity: 40.237

alignment_block:
US-09-486-167A-2 x AAC45603 ..

Align seg 1/1 to: AAC45603 from: 1 to: 674
```



```
*1 MetalaProfilLeuLysValGlyAspAlaIlePro.....AlaValGluVa 15
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
54 ATGGCTCCAAATTACTGTGCGGATGTTGTACACGACGGAACATCTCTTT 103
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
15 lphe...GluGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheL 31
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
104 CTTGATGAATAATGATCAGCTTCAGACCGTCTCTGTCTCACTCTATCGCCG 153
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
31 ysGlyLysLysGlyValLeuPheGlyValProGlyAlaPheThrProGly 47
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
154 CCGGTAAAAAGTCACTCTCTTGTGCTCTCTGCTGCTTCACTCCACA 203
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
48 CysSerLysThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLy 64
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
204 TCCAGCATGACCATGTCCTGATTCATTTGGAAAGCAGAGGAGCTGAA 253
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
64 salalysGlyValGlnValValAlaCysLeuSerValAsnAspAlaPheV 81
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
254 GTCAGAGGTATCGATGATATCTTCTTGTGATGATGATGATGATGATG 303
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
81 alThrGlyGluTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeu 97
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
304 TGATGAGGCATGGGAAAGACATACCAAGAGAACAGCATGTGAAGTTT 353
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
98 LeuAlaaspProThrGlyAlaPhe.....GlyLysGluThrAs 110
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
354 GTAGCAGATGGTGTGGAGATACACATCTTCTTGGACTTGAGCTTGA 403
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
110 pLeuLeuLeuAspSerLeuValSerIlePheGlyAsnArgArgLeuL 127
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
404 T.....CTCAGGACGAGGCTGTGTGAT.....AGGTCAA 435
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
127 ysArgPheSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
436 GGAGATTGCTCTCTGCTTGATTAACCTTAAGGTAACTGTAGCCATGTC 485
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
144 GluProaspGlyThrGlyLeuThrCysSerLeuAlaProAsnIleIleSe 160
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
486 GAA...AATGTTGCGAGTTTCACGGTTTCCAGCGCAGAGGATATCCCTAA 532
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
160 rGlnLeu 162
: |||
533 GGCTCTC 539

seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT.AAC45889
seq_documentation_block:
ID AAC45889 standard; DNA; 489 BP.
XX AAC45889;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 48136.
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 48136.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
```

```
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
```

```
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145214.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:
    Quality: 286.00      Length: 169
    Ratio: 2.509         Gaps: 6
Percent Similarity: 67.456 Percent Identity: 40.237

alignment_block:
US-09-486-167A-2 x AAC45889 ..

Align seg 1/1 to: AAC45889 from: 1 to: 489

1 MetAlaProIleLysValGlyAspAlaIlePro.....AlaValGluVa 15
|||||
1 ATGGCTCCAAATTACTGTGCGGATGTGTACACAGACGGAAGTCTCTTT 50

15 lPhe...GluGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheL 31
|||
51 CTTGATGAAATGATCAGCTTCAGACCGTCTCGTTCACTCTATCGCG 100

31 ysGlyLysLysGlyValLeuPheGlyValProGlyAlaPheThrProGly 47
|||||
101 CCGTAAAGGATGATCTCTTTGGTGTCTGTTTCTTCACTCCACCA 150

48 CysSerLysThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLy 64
|||||
151 TGCAGTATGAGCCATGTCCTGGATTCTATGGGAAACGAGAGAGCTGAA 200

64 sAlaLysGlyValGlnValValAlaCysLeuSerValAsnAspAlaPheV 81
|||||
201 GTCAAAGGTTATTGATGAGATCTTGTCTTAGTGTGAATGATCATTTG 250

81 alThrGlyGluTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeu 97
|||
251 TGATGAAGGATGGGAAACATATCCAGAGAAGCATGTGAAGTTT 300

98 LeuAlaAspProThrGlyAlaPhe.....GlyLysGluThrAs 110
|||||
301 GTACAGATGGTCTGGAGAAATACACGCATCTTCTTGACTTGAGCTTGA 350

110 pLeuLeuAspAspSerLeuValSerIlePheGlyAsnArgArgLeuL 127
|
351 C.....CTTAAGGACAAAGGTCCTTGTATT.....AGGTCAA 382

127 ysArgPheSerMetValValGlnAspGlyIleValLysAlaLeuAsnVal 143
:|||||
:|||||
```

```

343 GGAGATTCCCTTTGCTTGCTTGATAACCTTAAGGTGACTGTAGCCCAATGTT 432
144 GluProAspGlyThrGlyLeuThrCysSerLeuAlaProAsnIleIleSe 160
   |||      |||      |||      |||      |||      |||      |||      |||
433 GAATCTGGTGGCGAG...TTCACGGTTTCCAGCGCAGGAGGATATCTCAA 479
160 rGlnLeu 162
   :      |||
480 GGCTCTT 486

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2001, 11:20:20 ; Search time 109.78 Seconds
(without alignments)
4604.307 Million.cell updates/sec

Title: US-09-486-167A-1

Perfect score: 805

Sequence: 1 gccaggagcgagtggaag.....ttgtgtttgcggaaaaaaa 805

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601:*

- 1: /SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	805	100.0	805	20	AAAX27965
2	798	99.1	993	20	AAAX56411
3	462	57.4	604	20	AAAX27968
4	456	56.6	601	20	AAAX27966
5	408.6	50.8	453	21	AAAC04049
6	348.6	43.3	452	22	AAAF92316
7	300.8	37.4	469	20	AAAX27967
8	299	37.1	315	16	AAAT21145
9	209	26.0	423	21	AAAC02224
10	197	24.5	2710	20	AAAX27969
11	126.6	15.7	553	21	AAAC95054

ALIGNMENTS

RESULT 1

ID AAAX27965 standard; cDNA; 805 BP.

XX AC AAAX27965;

XX DT 08-JUN-1999 (first entry)

XX DE Human bronchoalveolar polypeptide, B18hum, coding sequence.

XX KW B18hum; bronchoalveolar protein; peroxisome-associated polypeptide; lung injury; oxidative stress-related disorder; inflammatory disease; KW cardiovascular disease; neurodegenerative disorder; allergic reaction; KW amyotrophic lateral sclerosis; apoptosis; high bone mass syndrome; KW osteopetrosis; osteoporosis-pseudoglioma syndrome; diagnosis; KW Bardet-Biedl syndrome 1; therapy; ss.

XX OS Homo sapiens.

XX PN WO9909054-A2.

XX PD 25-FEB-1999.

XX PF 20-AUG-1998; 98WO-BE00124.

XX PR 20-AUG-1997; 97BE-0000692.

XX PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.

XX PI (UYMO-) UNIV MONS-HAINAUT.

XX DR Bernard A, Falmagne P, Hermans C, Knoops B, Wattiez R;

XX WPI; 1999-180968/15.

XX P-PSDB; AAY01079.

12	74.6	9.3	674	21	AAAC45603	Arabidopsis thalia
13	74.6	9.3	678	21	AAC34954	Arabidopsis thalia
14	73.6	9.1	797	21	AAC33738	Arabidopsis thalia
15	71.4	8.9	737	21	AAC34962	Arabidopsis thalia
16	71	8.8	811	21	AAC35188	Arabidopsis thalia
17	69.8	8.7	473	21	AAC45698	Arabidopsis thalia
18	69.8	8.7	489	21	AAC45889	Arabidopsis thalia
19	69.8	8.7	683	21	AAC35640	Arabidopsis thalia
20	68.6	8.5	618	18	AAT85874	Malassezia fungus
21	68.2	8.5	727	21	AAC33030	Arabidopsis thalia
22	67.6	8.4	789	21	AAF12020	Aspergillus oryzae
23	65.2	8.1	551	18	AAT85875	Malassezia fungus
24	63.8	7.9	559	21	AAF11315	Aspergillus niger
25	60.2	7.5	58909	22	AAF28543	Genomic fragment #
26	59	7.3	812	21	AAC42111	Arabidopsis thalia
27	57	7.1	320	21	AAAC45600	Arabidopsis thalia
28	57	7.1	991	18	AAT85884	Malassezia fungus
29	55.2	6.9	705	21	AAF14598	Aspergillus oryzae
30	54.8	6.8	462	21	AAC43548	Zea mays DNA fragm
31	54.4	6.8	534720	19	AAV30458	Rhizobium species
32	54.4	6.8	536165	19	AAV30459	Rhizobium species
33	53	6.6	5556	21	AAAB1732	N. meningitidis pa
34	53	6.6	349980	21	AAF21609	Neisseria meningit
35	50.8	6.3	495	21	AAC43306	Arabidopsis thalia
36	44.2	5.5	600	18	AAT85883	Malassezia fungus
37	41	5.1	1337	20	AAZ17263	Human gene express
38	40.6	5.0	1218	21	AAAO2488	Human colon cancer
39	39.8	4.9	114955	20	AAAS3491	Human adenosine A1
40	38.4	4.9	4616	17	AAAT9604	HIC-1 gene. Homo
41	38.4	4.8	554	21	AAF08758	Fusarium venenatum
42	38.4	4.8	1014	21	AAF18187	Lung cancer associ
43	38	4.7	1000	21	AAAO2484	Human colon cancer
44	37.4	4.6	1920	21	AAAC46281	Arabidopsis thalia
45	37.4	4.6	2090	21	AAAC45029	Arabidopsis thalia

Qy 541 ctggtgtccatcttttgggaatcgacgtcttcaagaggttctccatggtggtacaggaatgac 600

Db 346 ggaacagacttattactagatgattcgctgggtgtccatcttttgggaatcgacgtctcaa 405
QY 573 gagttctccatggtgtacagatgcatagtgaaaggccctgaattgtgaaacagatgg 632
Db 406 gagttctccatggtgtacagatgcatagtgaaaggccctgaattgtgaaacagatgg 465
QY 633 cacaggcctcaactgcaagctggcaccacaatacatctacacgtcttgaggccctggcc 692
Db 466 cacaggcctcaactgcaagctggcaccacaatacatctacacgtcttgaggccctggcc 525
QY 693 agattacttctccacccctccctatctccctgcccagccctgtgtgaggccctgcaa 752
Db 526 agattacttctccacccctccctatctccctgcccagccctgtgtgaggccctgcaa 585
QY 753 ttggaattgtggccag 768
Db 586 ttggaattgtggccag 601

RESULT 5
AAC04049
ID AAC04049 standard; cDNA; 453 BP.
AC AAC04049;
XX
XX 06-OCT-2000 (first entry)
DT
DE Human secreted protein 5' EST, SEQ ID NO: 4047.
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX P-PSDB; AAG04043.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 4047; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX mRNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors.
XX
XX Sequence 453 BP; 82 A; 108 C; 180 G; 82 T; 1 other;

Query Match 50.8%; Score 408.6; DB 21; Length 453;
Best Local Similarity 99.5%; Pred. No. 3.2e-103;
Matches 419; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 gccagagagcgcagtggaagtggccgtggcgccgtatgggaactagctggcgtgtgcgcc 60
Db 34 gccagagagcgcagtggaagtggccgtggcgccgtatgggaactagctggcgtgtgcgcc 93
QY 61 ctgagacgtctcagcgggtctatactcgtggtggcgccggtcagctctcgcgagcagc 120
Db 94 ctgagacgtctcagcgggtctatactcgtggtggcgccggtcagctctcgcgagcagc 153
QY 121 gcagcaagacgtctcagtgaaaggagagtggtggtggtggtggtggtggtggtggtggt 180
Db 154 gcagcaagacgtctcagtgaaaggagagtggtggtggtggtggtggtggtggtggtggt 213
QY 181 gccgctgcagccatgcccccaatcaagtggtggtggtggtggtggtggtggtggtggt 240
Db 214 gccgctgcagccatgcccccaatcaagtggtggtggtggtggtggtggtggtggtggt 273
QY 241 gaaggggagccagggaaacaaagtgaaacctggcagagctgttcaagggcaagaagggtgtg 300
Db 274 gaaggggagccagggaaacaaagtgaaacctggcagagctgttcaagggcaagaagggtgtg 333
QY 301 ctgtttggagttctcgtgggcttccacctggtatgttccaaagacacacacacacacacac 360
Db 334 ctgtttggagttctcgtgggcttccacctggtatgttccaaagacacacacacacacacac 393
QY 361 gtggagcagctgagcctctgaaggccaaggaggtggtggtggtggtggtggtggtggtggt 420
Db 394 gtggagcagctgagcctctgaaggccaaggaggtggtggtggtggtggtggtggtggtggt 452
QY 421 a 421
Db 453 a 453

RESULT 6
AAF92316
ID AAF92316 standard; cDNA; 452 BP.
XX
XX AAF92316;
AC
XX
DT 15-MAY-2001 (first entry)
XX
XX Bovine mammary tissue derived cDNA #29.
XX
XX Bovine; mammary gland; cancer; tumour; angiogenesis; ss.
XX
XX Bos taurus.
OS
XX
XX WO200114553-A1.
PN
XX
XX 01-MAR-2001.
PD
XX
XX 23-AUG-2000; 2000WO-NZ00166.
PF
XX
XX 23-AUG-1999; 99US-0150330.
PR
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA
XX (NZPA-) NEW ZEALAND PASTORAL AGRIC RES INST LTD.
XX
XX Havukkala IJ, Gleen M, Grigor MR, Molenaar AJ;
XX
XX WPI; 2001-226619/23.
DR
XX
XX New polypeptides and polynucleotides encoding the polypeptides, which
XX are expressed in bovine mammary gland tissue, useful for stimulating
XX mammary gland growth or function, or inducing differentiation of milk
XX producing cells
XX
XX Claim 1; Page 49-50; 97pp; English.
XX

Db 426 tgccacgccgtgtgctgtggccctgcaattggaattgttgccag 469

RESULT 8

AAT21145

ID AAT21145 standard; cDNA to mRNA; 315 BP.

XX

AC AAT21145;

XX

DT 01-AUG-1996 (first entry)

XX

DE Human gene signature HUMGS02452.

XX

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency; human; cloning; mapping; non-biased library; diagnosis; detection; cell typing; abnormal cell function; ss.

XX

OS Homo sapiens.

XX

PN W09514772-A1.

XX

PD 01-JUN-1995.

XX

PF 11-NOV-1994; 94WO-JP01916.

XX

PR 12-NOV-1993; 93JP-0355504.

XX

PA (MATS/) MATSUBARA K.

XX

PA (OKUB/) OKUBO K.

XX

PI Matsubara K, Okubo K;

XX

PN WPI; 1995-206931/27.

XX

PT Identifying gene signatures in 3'-directed human cDNA library - e.g. for diagnosis of abnormal cell function, by preparing cDNA that reflects relative abundance of corresp. mRNA in specific human tissues

PT

XX

PS Claim 1; Page 813; 2245pp; Japanese.

XX

CC A single-stranded DNA (or its complementary strand or the corresp. double-stranded DNA) which comprises one of the 7837 "GS" sequences given in AAT19001-T26837 and which is able to hybridise to part of human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were obtained from 3'-directed cDNA libraries prepared from various human tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-untranslated sequence is unique to a particular mRNA species, almost all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library is constructed so as to reflect accurately the relative abundance of different mRNAs in the particular tissue from which it was derived. The appearance frequency of a given GS in a cDNA library can be determined (esp. using primers and probes derived from the GS sequences) as a means of diagnosing abnormal cell function or for recognising different cell types.

XX

Sequence 315 BP; 69 A; 82 C; 81 G; 75 T; 8 other;

Query Match

Best Local Similarity 37.1%; Score 299; DB 16; Length 315;

Matches 302; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 490 gatccacatggcccttggaggagacagacttattactagatgattcgtggtgcc 549

Db 1 gatccacatggggccttggaggagacagacttattactagatgattcgtggtgcc 60

QY 550 attcttgggaatgcagctctcaagagttctccatggtgtgacagatgagtgag 609

Db 61 attcttgggaatgcagctctcaagagttctccatggtgtgacagatgagtgag 120

QY 610 ggcctgaattggaacagatggcacagcgcctcaactgcagcctgcccacccaatc 669

Db 121 gccctgaattggaacagatggcacagcgcctcaactgcagcctgcccacccaatc 180

QY 670 tcacagctctgagggccctggccagattacttctccacccctccctatctacactgcc 729

Db 181 tcacagctctgagggccctggccagattacttctccacccctccctatctacactgcc 240

QY 730 agccctgtgctggggccctggccagattggaattgttggccagatttctgcaataaacctgt 789

Db 241 agccctgtgctggggccctggccagattggaattgttggccagatttctgcaataaacctgt 300

QY 790 ggtttgcggaaaaa 804

Db 301 ggtttgcgggtcaaa 315

RESULT 9

AAC02224

ID AAC02224 standard; cDNA; 423 BP.

XX

AC AAC02224;

XX

DT 06-OCT-2000 (first entry)

XX

DE Human secreted protein 5' EST, SEQ ID NO: 2222.

XX

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping; ss.

XX

OS Homo sapiens.

XX

PN EP1033401-A2.

XX

PD 06-SEP-2000.

XX

PF 21-FEB-2000; 2000EP-0200610.

XX

PR 26-FEB-1999; 99US-0122487.

XX

PA (GEST) GENSET.

XX

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX

WPI; 2000-500381/45.

DR P-PSDB; AAG02218.

XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for

PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for

PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX

PS Claim 1; SEQ ID 2222; 71pp + CD-ROM; English.

XX

CC The present sequence is one of a large number of 5' ESTs derived from

CC mRNAs encoding secreted proteins. An ORF has been identified within the

CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs

CC derived from 30 different tissues. EST sequences usually correspond

CC mainly to the 3' untranslated region (UTR) of the mRNA because they are

CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not

CC well suited for isolating cDNA sequences derived from the 5' ends of

CC mRNAs and even in those cases where longer cDNA sequences have been

CC obtained the full 5' UTR is rarely included. 5' ESTs are derived from

CC mRNAs with intact 5' ends and can therefore be used to obtain full length

CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,

CC gene therapy and chromosome mapping procedures. They are used to obtain

CC upstream regulatory sequences and to design expression and secretion

CC vectors.

XX

Sequence 423 BP; 77 A; 109 C; 156 G; 81 T; 0 other;

QY

Query Match

Best Local Similarity 26.0%; Score 209; DB 21; Length 423;

Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

KW	metabolic pathway; promoter; termination sequence; ss
XX	
XX	Arabidopsis thaliana.
OS	
XX	EP1033405-A2.
PN	
XX	
XX	06-SEP-2000.
PD	
XX	
XX	25-FEB-2000; 2000EP-0301439.
PF	
XX	
XX	25-FEB-1999; 99US-0121825.
PR	
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130416.

[illegible]

18-JUN-1999	PR	9905-0139458
18-JUN-1999	PR	9905-0139459
18-JUN-1999	PR	9905-0139460
18-JUN-1999	PR	9905-0139461
18-JUN-1999	PR	9905-0139462
18-JUN-1999	PR	9905-0139463
18-JUN-1999	PR	9905-0139750
18-JUN-1999	PR	9905-0139763
21-JUN-1999	PR	9905-0139817
22-JUN-1999	PR	9905-0139899
23-JUN-1999	PR	9905-0140353
23-JUN-1999	PR	9905-0140354
24-JUN-1999	PR	9905-0140695

PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147024.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149773.			
PR	20-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151066.			
PR	30-AUG-1999;	99US-0151303.			
PR	31-AUG-1999;	99US-0151438.			
PR	01-SEP-1999;	99US-0151930.			
PR	07-SEP-1999;	99US-0152363.			
Query Match 9.3%; Score 74.6; DB 21; Length 674;					
Best Local Similarity 59.0%; Pred. No. 3.8e-11;					
Matches 128; Conservative 0; Mismatches 89; Indels 0; Gaps 0;					
Qy	286	ggcaagaagggtgtgtctgttttgaggattctctggggccttcaccccttgatgttccaaagaca	345		
Db	156	ggtaaaaaagtcattctcttctgtgttctctgtgttctctcactccacatgcagcatgagc	215		
Qy	346	cacctgccagggtttgtgagcaggctgagctctgaaggcccaaggaggtccagtggtg	405		
Db	216	catgtgcctggattcattggaaagcagagagctgaaggtcaagggttatcgatgagatc	275		
Qy	406	gcctgtctgagtgttaatgatgcctttgtgactggcggagtgggccgagcccaaggcg	465		
Db	276	attgtcttagtgtgaatgatccatttgtgatgaaggcatggggaagacataccaagag	335		
Qy	466	gaaggcaagggttcgctcctggtgagtcacccactggggg	502		
Db	336	aacaagcatgtgaadttgtgacagatgggtctgag	372		
RESULT 13					
AAC34954					
ID	AAC34954 standard; DNA; 678 BP.				
XX	AAC34954;				
AC	AAC34954;				
XX	AAC34954;				
DT	17-OCT-2000 (first entry)				

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 8494.
DE XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
PD XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134236.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151338.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	28-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161922.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match		
Best Local Similarity 59.3%; Score 74.6; DB 21; Length 678;		
Matches 128; Conservative 0; Mismatches 89; Indels 0; Gaps		
QY	286	ggcaagaaaggttgctgtgttgagttcctgggcttcacccctggatgttccaagaca 345
Db	157	ggtaaaaaagtcattctcttgggttctcgtgtgtttcaactccacatgaagcatgac 216
QY	346	cacctgcacaggtttgtgagacagctagagctcgaaggccaaggaggtccagtggtg 405
Db	217	catgtgcctggattccatgggaaagcadagggagctgaagtcgaaggttatcgatgagatc 276
QY	406	gctgtctgagtgttaatgatgcctttgtactggcagtgaggccagagcccaagggc 465
Db	277	attgtcttagtgtgaatgatccattgtgatgaaggcatgggggaaaaacacacagag 336
QY	466	gaaggcaaggttcgctcctcgctgatccactgggg 502
Db	337	aacagcatgtgaagtttgtacagatggatctggag 373

ID	AAC33738	standard; DNA; 797 BP.
XX	AAC33738;	
AC	17-OCT-2000	(first entry)
XX	Arabidopsis thaliana	DNA fragment SEQ ID NO: 4141.
DE	Hybridisation assay; genetic mapping; gene expression control;	
XX	protein identification; signal transduction pathway;	
KK	metabolic pathway; promoter; termination sequence; ss.	
KK	Arabidopsis thaliana.	
OS	EP1033405-A2.	
XX	06-SEP-2000.	
PN	25-FEB-2000;	2000EP-0301439.
XX	25-FEB-1999;	99US-0121825.
PPR	05-MAR-1999;	99US-0123180.
PPR	09-MAR-1999;	99US-0123548.
PPR	23-MAR-1999;	99US-0125788.
PPR	25-MAR-1999;	99US-0126264.
PPR	29-MAR-1999;	99US-0126785.
PPR	01-APR-1999;	99US-0127462.
PPR	06-APR-1999;	99US-0128234.
PPR	08-APR-1999;	99US-0128714.
PPR	16-APR-1999;	99US-0129845.
PPR	19-APR-1999;	99US-0130077.
PPR	21-APR-1999;	99US-0130449.
PPR	23-APR-1999;	99US-0130510.
PPR	23-APR-1999;	99US-0130891.
PPR	28-APR-1999;	99US-0131449.
PPR	30-APR-1999;	99US-0132048.
PPR	30-APR-1999;	99US-0132407.
PPR	04-MAY-1999;	99US-0132484.
PPR	05-MAY-1999;	99US-0132485.
PPR	06-MAY-1999;	99US-0132486.
PPR	06-MAY-1999;	99US-0132487.
PPR	07-MAY-1999;	99US-0132863.
PPR	11-MAY-1999;	99US-0134256.
PPR	14-MAY-1999;	99US-0134218.
PPR	14-MAY-1999;	99US-0134219.
PPR	14-MAY-1999;	99US-0134221.
PPR	14-MAY-1999;	99US-0134370.
PPR	18-MAY-1999;	99US-0134768.
PPR	19-MAY-1999;	99US-0134941.
PPR	20-MAY-1999;	99US-0135124.
PPR	21-MAY-1999;	99US-0135353.
PPR	24-MAY-1999;	99US-0135629.
PPR	25-MAY-1999;	99US-0136021.
PPR	27-MAY-1999;	99US-0136392.
PPR	28-MAY-1999;	99US-0136782.
PPR	01-JUN-1999;	99US-0137222.
PPR	03-JUN-1999;	99US-0137528.
PPR	04-JUN-1999;	99US-0137502.
PPR	07-JUN-1999;	99US-0137724.
PPR	08-JUN-1999;	99US-0138094.
PPR	10-JUN-1999;	99US-0138540.
PPR	10-JUN-1999;	99US-0138847.
PPR	14-JUN-1999;	99US-0139119.
PPR	16-JUN-1999;	99US-0139452.
PPR	16-JUN-1999;	99US-0139453.
PPR	17-JUN-1999;	99US-0139492.
PPR	18-JUN-1999;	99US-0139454.
PPR	18-JUN-1999;	99US-0139455.
PPR	18-JUN-1999;	99US-0139456.
PPR	18-JUN-1999;	99US-0139457.
PPR	18-JUN-1999;	99US-0139458.
PPR	18-JUN-1999;	99US-0139459.
PPR	18-JUN-1999;	99US-0139460.

Db 447 acaaacatgtgaagttttagctgtagtggtgctggag 482

RESULT 15

AAC34962

ID AAC34962 standard; DNA; 737 BP.

XX AAC34962;

AC AAC34962;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 8523.

DE Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

OS EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126284.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135622.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 03-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_pr9: HSA249483	+	830.00	1418.11	9.4e-71	489	! AJ249483 Homo sapiens mRNA for
gb_pr5: AF231705	+	830.00	1414.28	1.5e-70	788	! AF231705 Homo sapiens Alu con-
gb_pat1: A99097	+	830.00	1414.11	1.6e-70	805	! A99097 Sequence 1 from Patient W
gb_pr4: AF110731	+	830.00	1414.11	1.6e-70	805	! AF110731 Homo sapiens antioxidant
gb_pr5: AF112212	+	830.00	1413.82	1.6e-70	835	! AF112212 Homo sapiens putative
gb_pr5: AF112493	+	830.00	1413.77	1.6e-70	845	! AF112493 Homo sapiens peroxisom
gb_pr5: AF242525	+	830.00	1413.44	1.7e-70	875	! AF242525 Homo sapiens peroxisom
gb_pr4: AF110736	+	822.00	1401.35	8.2e-70	845	! AF110736 Cercopithecus aethiops
gb_pr5: AF197952	+	820.00	1401.25	1.4e-69	780	! AF197952 Homo sapiens thioether
gb_pr4: AF110734	+	818.00	1397.40	2.0e-69	725	! AF110734 Papio hamadryas peroxi
gb_pat1: A99101	+	767.00	1397.66	1.3e-64	675	! A99101 Sequence 5 from Patient W
gb_pr4: AF110733	+	767.00	1397.66	1.3e-64	675	! AF110733 Mus musculus peroxired
gb_r01: AF197951	+	767.00	1307.51	1.4e-64	688	! AF197951 Mus musculus thioether
gb_r01: AF110732	+	767.00	1305.93	1.7e-64	837	! AF110732 Rattus norvegicus anti
gb_r01: AF124994	+	767.00	1305.40	1.7e-64	895	! AF124994 Mus musculus peroxisom
gb_pat1: AF110735	+	761.00	1296.04	5.9e-64	798	! AF110735 Sus scrofa peroxiredox
gb_pat1: A99099	+	752.00	1280.82	4.2e-63	881	! A99099 Sequence 3 from Patient W
gb_r01: AF305564	+	747.00	1271.28	1.4e-62	780	! AF305564 Bos taurus peroxiredox
gb_pat1: A99104	+	597.50	1009.29	5.6e-48	1869	! A99104 Sequence 8 from Patient W
gb_r01: A99105	+	568.00	976.44	3.8e-46	601	! A99105 Sequence 9 from Patient W
gb_r01: AF208729S2	+	498.50	837.92	1.1e-45	2360	! AF208730 Homo sapiens peroxire
gb_htg3: AP001453	+	486.00	783.13	2.2e-35	151747	! AP001453 Homo sapiens chromo
gb_htg3: AC0015698	+	486.00	782.29	2.5e-35	168475	! AC0015698 Homo sapiens chromo
gb_htg1: AC0005848	+	486.00	782.00	2.5e-35	174562	! AC005848 Homo sapiens chromo
gb_pat1: A99106	+	476.50	799.14	2.8e-36	2710	! A99106 Sequence 10 from Patient
gb_pr10: HSU82616	+	463.50	789.40	9.8e-36	569	! U82616 Human mRNA downregulat
gb_htg3: AC001380	+	448.00	717.74	9.7e-32	157974	! AC001380 Homo sapiens chromo
gb_htg3: AC014470	+	434.50	690.25	3.3e-30	33270	! AC014470 Drosophila melanoga
gb_in1: AC009462	-	424.50	677.84	1.6e-29	151611	! AC009462 Drosophila melanoga
gb_in2: AE003721	-	424.50	674.68	2.4e-29	224896	! AE003721 Drosophila melanoga
gb_htg12: AC0449118	+	417.00	662.97	1.1e-28	195316	! AC044918 Homo sapiens chromo
gb_htg19: AL139155	+	406.00	647.57	7.8e-28	127234	! AL139155 Homo sapiens chromo
gb_htg23: AC000573	+	395.00	633.97	4.5e-27	66266	! AP000573 Homo sapiens chromo
gb_tst1: CN0007C9C	+	378.50	639.52	6.2e-27	979	! AL438886 T7 end of clone BD0AA
gb_pat1: CN007C9K	+	373.50	631.00	6.6e-27	974	! AL439722 T7 end of clone BD0AA
gb_pat1: A99103	+	349.00	594.91	6.7e-25	469	! A99103 Sequence 7 from Patient W
gb_htg18: AC0087899_3	+	345.00	544.81	4.1e-22	103152	! Continuation (4 of 4) of AC0
gb_p13: CN001DPB	+	325.50	552.12	1.6e-22	645	! AL117111 Botrytis cinerea stra
gb_p13: CN001CP2	+	321.50	545.85	3.6e-22	600	! AL115806 Botrytis cinerea stra
gb_p13: CN001CVS	+	321.50	544.39	4.4e-22	720	! AL116048 Botrytis cinerea stra

Mon Oct 22 14:46:56 2001

us-09-486-167a-2.rge

```

151 ACACACCTGCCAGGCTTTGTGAGCAGGCTGAGGCTCTGAAGGCCAAGG 200
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
201 AGTCCAGGTGGTGGCTGCTGAGTGTAAATGATGCTTTGTGACTGGCG 250
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
251 AGTGGGCCGAGCCCAACAGCGGAGGCAAGGTTCCGGCTCTCGGTGAT 300
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
301 CCCACTGGGCGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 350
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
351 GGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTTCTCCATGGTGTAC 400
134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
401 AGGATGGCAGTAGTGAAGGCCCTCAATGTGAACACAGATGGCACAGGCCTC 450
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
451 ACCTGCAGCTGGCAGCCCAATATCATCTCACAGCTC 486

```

seq_name: gb_pr5:AF231705

seq_documentation_block: 788 bp mRNA PRI 27-JUN-2000
LOCUS AF231705 Homo sapiens Alu co-repressor 1 (ACR1) mRNA, complete cds.

ACCESSION AF231705
VERSION AF231705.1 GI:8745393

KEYWORDS human.

SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 788)
Kropotov,A., Sedova,V., Ivanov,V., Sazeeva,N., Tomilin,A.,
Krutillina,R., Oei,S.L., Griesenbeck,J., Buchlow,G. and Tomilin,N.
A novel human DNA-binding protein with sequence similarity to a
subfamily of redox proteins which is able to repress
RNA-polymerase-III-driven transcription of the Alu-family
retroposons in vitro
Eur. J. Biochem. 260 (2), 336-346 (1999)

JOURNAL MEDLINE 99195471
PUBMED 10095767

REFERENCE 2 (bases 1 to 788)
Kropotov,A.V. and Tomilin,N.V.
Direct Submission
Submitted (04-FEB-2000) Chromosome Stability, Institute of Cytology
RAS, Tikhoretskii Av.4, St.Petersburg 194064, Russia

FEATURES
source Location/Qualifiers

```

1..788
/organism="Homo sapiens"
/db_xref="taxon.9606"
/chromosome="11"
/map="11q13"
1..788
/gene="ACR1"
20..664
/gene="ACR1"
/codon_start=1
/product="Alu co-repressor 1"
/protein_id="AAF78899.1"
/db_xref="GI:8745394"
/translation="MGLAGVCAIARRSAGYILVGGAGGQSAARARRCSEGEWASGGVR
SFSPAAAMAPIKVGDAIPAVEVEPEGPNKNLAEFLKKGKGVLFVGPAGFTPGCSK
THLPFGVEQAEAKAGQVQVACISVNDFAVTGEMGRAHKAEGKVRLLADPTGAFGRE
TDLIDSLVSIFGNRRILKRFSEWVQDGIYKALNVPEPDGTLTCSLAPNIISQL"

```

gene

CDS

BASE COUNT 153 a 201 c 268 g 166 t

ORIGIN

```

alignment_scores:
Quality: 830.00 Length: 162
Ratio: 5.123 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-486-167A-2 x AF231705 ..
Align seg 1/1 to: AF231705 from: 1 to: 788

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG 17
176 ATGGCCCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 225
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
226 AGGGAGCCAGGGAACCAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 275
34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
276 AGGGTGTGCTGTTGGAGTTCCTGGGGCCTTCACCCCTGGATGTTCCAAG 325
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG 67
326 ACACACCTGCCAGGCTTTGTGAGCAGGCTGAGGCTCTGAAGGCCAAGG 375
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
376 AGTCCAGGTGGTGGCCTGTCTGAGTGTAAATGATGCTTTGTGACTGGCG 425
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
426 AGTGGGCCGAGCCCAACAGCGGAGGCAAGGTTCCGGCTCTCGGTGAT 475
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
476 CCCACTGGGCGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCCT 525
117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
526 GGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTTCTCCATGGTGTAC 575
134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
576 AGGATGGCAGTAGTGAAGGCCCTCAATGTGAACACAGATGGCACAGGCCTC 625
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
626 ACCTGCAGCTGGCAGCCCAATATCATCTCACAGCTC 661

```

seq_name: gb_pat1:A99097

seq_documentation_block: 805 bp DNA PAT 20-SEP-2000
LOCUS A99097
DEFINITION Sequence 1 from Patent WO9909054.
ACCESSION A99097
VERSION A99097.1 GI:6782047

KEYWORDS human.

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

Patent: WO 9909054-A 1 25-FEB-1999;
UNIV MONS HAINAUT (BE) ; FALMAGNE PAUL (BE) ; WATTIEZ RUDDY (BE) ;

```

*
BERNARD ALFRED (BE) ; HERMANS CEDRIC (BE) ; KNOOPS BERNARD (BE) ;
UNIV LOUVAIN (BE)
FEATURES
  source
    1..805
    /location=Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    193..681
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAB70589.1"
    /db_xref="GI:6782048"
    /translation="MAPIRKVGDAIPAVEFEGEPGNKNVLAELFRGKGVLFVPGAF
TPGSKTHLPFGVEQAEALKAQGVVACLVSNDAFVTGEMGRAHKAQGVKRLADPT
GAPGKETDLDLSDLSVIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNIS
QL"
  CDS
    163 a 200 c 276 g 166 t
  alignment_scores:
    Quality: 830.00 Length: 162
    Ratio: 5.123 Gaps: 0
    Percent Similarity: 100.000 Percent Identity: 100.000
  alignment_block:
    US-09-486-167A-2 x A99097 ..
  Align seg 1/1 to: A99097 from: 1 to: 805
    1 MetAlaProLLeysValGlyAspAlaIleProAlaValGluValPheG1 17
    193 ATGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 242
    17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34
    243 AGGGACCCAGGGAACAAAGTGAACCTGGCAGAGCTGTTCAAGGCAAGA 292
    34 ySGLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
    293 AGGTGTGCTGTTGGAGTCTCTGGGGCTTCACCCCTGGATGTTCCAAG 342
    51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
    343 ACACACTGCCAGGGTTGTGGAGCAGGCTGAGGCTCTCAAGGCCAAGG 392
    67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
    393 AGTCCAGGTGGTCTGCTGAGTGTATGATGCTTTGTGACTGGCG 442
    84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
    443 AGTGGGGCCGACCCACAAAGCGGAAGCAAGGTTCCGCTCTGCTGAT 492
    101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
    493 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATCGCT 542
    117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
    543 GGTGTCCATCTTTGGGAATCAGAGCTCTCAAGAGGTTCTCCATGGTGTAC 592
    134 InAspGlyLeValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
    593 AGGATGGCATAGTGAAGCCCTGATGTGGAACCAAGATGGCACAGGCCTC 642
    151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
    643 ACCTGCAGCTGGCACCAATATCATCTCACAGCTC 678
  seq_name: gb_pr4:AF110731
  seq_documentation_block:
  LOCUS
    AF110731 805 bp mRNA PRI 26-SEP-2000

```

DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM

Homo sapiens antioxidant enzyme B166 mRNA, complete cds.
AF110731
AF110731.1 GI:6103723
human.
Homo sapiens

REFERENCE AUTHORS TITLE

1 (bases 1 to 805)
Wattiez,R., Hermans C., Bernard,A., Lesur,O. and Falmagne,P.
Human bronchoalveolar lavage fluid: two-dimensional gel
electrophoresis, amino acid microsequencing and identification of
major proteins
Electrophoresis 20 (7), 1634-1645 (1999)
99351776
10424490
2 (bases 1 to 805)
Knoops,B., Clippe,A., Bogard,C., Arsalane,K., Wattiez,R.,
Hermans,C., Duconseille,E., Falmagne,P. and Bernard,A.
Cloning and characterization of AOEB166, a novel mammalian
antioxidant enzyme of the peroxiredoxin family
J. Biol. Chem. 274 (43), 30451-30458 (1999)
99452929
10521424
3 (bases 1 to 805)
Knoops,B., Wattiez,R., Falmagne,P., Hermans,C. and Bernard,A.
Direct Submission
Submitted (04-DEC-1998) Department of Biology, Catholic University
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
Location/Qualifiers
1..805
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="11q13; about 7cR from marker D11S913; between
markers D11S1963 and D11S4407"
/tissue_type="lung"
37..681
/note="hAOEB166; alkyl hydroperoxide reductase;
thioredoxin peroxidase; represents prxv peroxiredoxin
subfamily in mammals"
/codon_start=1
/product="antioxidant enzyme B166"
/protein_id="AAF03750.1"
/db_xref="GI:6103724"
/translation="MGLAGVCALRRSAGYILVGGAGGSAASAAARRCSEGEWASGGVR
SFSRAAAMAPIKVGDAPVAVEFEGEPGNKNVLAELFRGKGVLFVPGAFPGCSK
THLPFGVEQAEALKAQGVVACLVSNDAFVTGEMGRAHKAQGVKRLADPTGAFGKE
TDLLDLSLSVIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNISQL"
193

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE

TITLE

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

FEATURES source

source

CDS

CDS

misc_feature

misc_feature

BASE COUNT ORIGIN

BASE COUNT 163 a 200 c 276 g 166 t
ORIGIN

alignment_scores:

alignment_scores:

Percent Similarity:

Percent Similarity:

alignment_block:

alignment_block:

US-09-486-167A-2 x AF110731

US-09-486-167A-2 x AF110731

Align seg 1/1 to: AF110731 from: 1 to: 805

Align seg 1/1 to: AF110731 from: 1 to: 805

1 MetAlaProLLeysValGlyAspAlaIleProAlaValGluValPheG1 17

1 MetAlaProLLeysValGlyAspAlaIleProAlaValGluValPheG1 17

193 ATGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 242

193 ATGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 242

17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34

17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34

243 AGGGACCCAGGGAACAAAGTGAACCTGGCAGAGCTGTTCAAGGCAAGA 292

243 AGGGACCCAGGGAACAAAGTGAACCTGGCAGAGCTGTTCAAGGCAAGA 292

34 ySGLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50

34 ySGLyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50

293 AGGTGTGCTGTTGGAGTCTCTGGGGCTTCACCCCTGGATGTTCCAAG 342

293 AGGTGTGCTGTTGGAGTCTCTGGGGCTTCACCCCTGGATGTTCCAAG 342

51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67

51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67

343 ACACACTGCCAGGGTTGTGGAGCAGGCTGAGGCTCTCAAGGCCAAGG 392

343 ACACACTGCCAGGGTTGTGGAGCAGGCTGAGGCTCTCAAGGCCAAGG 392

67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84

67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84

393 AGTCCAGGTGGTCTGCTGAGTGTATGATGCTTTGTGACTGGCG 442

393 AGTCCAGGTGGTCTGCTGAGTGTATGATGCTTTGTGACTGGCG 442

84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100

84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100

443 AGTGGGGCCGACCCACAAAGCGGAAGCAAGGTTCCGCTCTGCTGAT 492

443 AGTGGGGCCGACCCACAAAGCGGAAGCAAGGTTCCGCTCTGCTGAT 492

101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117

101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117

493 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATCGCT 542

493 CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATCGCT 542

117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134

117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134

543 GGTGTCCATCTTTGGGAATCAGAGCTCTCAAGAGGTTCTCCATGGTGTAC 592

543 GGTGTCCATCTTTGGGAATCAGAGCTCTCAAGAGGTTCTCCATGGTGTAC 592

134 InAspGlyLeValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150

134 InAspGlyLeValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150

593 AGGATGGCATAGTGAAGCCCTGATGTGGAACCAAGATGGCACAGGCCTC 642

593 AGGATGGCATAGTGAAGCCCTGATGTGGAACCAAGATGGCACAGGCCTC 642

151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162

151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162

643 ACCTGCAGCTGGCACCAATATCATCTCACAGCTC 678

643 ACCTGCAGCTGGCACCAATATCATCTCACAGCTC 678

seq_name: gb_pr4:AF110731

seq_name: gb_pr4:AF110731

seq_documentation_block:

seq_documentation_block:

LOCUS

LOCUS

AF110731

AF110731

805 bp

805 bp

mRNA

mRNA

PRI

PRI

26-SEP-2000

26-SEP-2000

163 a 200 c 276 g 166 t

163 a 200 c 276 g 166 t

Quality: 830.00

Quality: 830.00

Ratio: 5.123

Ratio: 5.123

Percent Similarity: 100.000

Percent Similarity: 100.000

Percent Identity: 100.000

Percent Identity: 100.000

alignment_block:

alignment_block:

US-09-486-167A-2 x AF110731

US-09-486-167A-2 x AF110731

Align seg 1/1 to: AF110731 from: 1 to: 805

Align seg 1/1 to: AF110731 from: 1 to: 805

1 MetAlaProLLeysValGlyAspAlaIleProAlaValGluValPheG1 17

1 MetAlaProLLeysValGlyAspAlaIleProAlaValGluValPheG1 17

193 ATGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 242

193 ATGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 242

17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34

17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34

243 AGGGACCCAGGGAACAAAGTGAACCTGGCAGAGCTGTTCAAGGCAAGA 292

243 AGGGACCCAGGGAACAAAGTGAACCTGGCAGAGCTGTTCAAGGCAAGA 292

Mon Oct 22 14:46:56 2001

us-09-486-167a-2.rge

BASE COUNT 198 a 199 c 272 g 166 t
ORIGIN

alignment_scores:
Quality: 830.00 Length: 162
Ratio: 5.123 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-486-167A-2 x AF112212 ..

Align seg 1/1 to: AF112212 from: 1 to: 835

1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
184 ATGGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTTGA 233
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34
234 AGGGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGCAAGA 283
34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
284 AGGGTGTGCTGTTGGAGTTCCTGGGGCCTTCACCCCTGGATGTTCCAAG 333
51 ThrHisLeuProGlyPheValGluAlaGluAlaLeuLysAlaLysG1 67
334 ACACACTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGGG 383
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
384 AGTCCAGGTGTGCTGCTGTGAGTGTAAATGATGCTTTGTGACGCGC 433
84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaA 100
434 AGTGGGCCGAGCCACAAAGCGGAAGGAGGTTTCGGCTCCTGGCTGAT 483
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117
484 CCCACTGGGGCCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 533
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
534 GGTGTCCTCTTTGGGAATCGACGCTCAAGAGGTTCTCCATGTTGTGAC 583
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
584 AGGATGCGATAGTGAAGGCCCTGAATGTGAACAGATGGCAGGCGCTC 633
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
634 ACCTGCAGCTGGGCCACCCATATATCATCTCACAGCTC 669

seq_name: gb_pr5:AF1124993

seq_documentation_block:
LOCUS AF124993 840 bp mRNA PRI 25-JAN-2000
DEFINITION Homo sapiens peroxisomal membrane protein 20 mRNA, complete cds.
ACCESSION AF124993
VERSION AF124993.1 GI:6746354
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Yamashita, H., Avraham, S., Jiang, S., London, R., Van Veldhoven, P.,
1 (bases 1 to 840)
TITLE Characterization of human and murine PMP20 peroxisomal proteins
JOURNAL that exhibit antioxidant activity in vitro
MEDLINE J. Biol. Chem. 274 (42), 29897-29904 (1999)
99445545

34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
293 AGGTGTGCTCTTGGAGTTCCTGGGGCTTCACCCCTGGATGTTCCAAG 342
51 ThrHisLeuProGlyPheValGluAlaGluAlaLeuLysAlaLysG1 67
343 ACACACCTGCCAGGGTTTGTGGAGAGGCTGAGGCTCTGAAGGCCAAGGG 392
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
393 AGTCCAGGTGGTGGCTGCTGAGTGTAAATGATGCTTCCTGGCTGATG 442
84 LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaA 100
443 AGTGGGCCGAGCCACAAAGCGGAAGCAAGGTTCGGCTCCTGGCTGAT 492
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117
493 CCCACTGGGGCCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 542
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
543 GGTGTCCTCTTTGGGAATCGAGCTCTCAAGAGGTTCTCCATGTTGGTAC 592
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
593 AGGATGCGATAGTGAAGGCCCTGAATGTGAACAGATGGCAGAGGCCCTC 642
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
643 ACCTGCAGCTGGGCCACCCATATCATCTCACAGCTC 678

seq_name: gb_pr5:AF112212

seq_documentation_block:
LOCUS AF112212 835 bp mRNA PRI 13-DEC-1999
DEFINITION Homo sapiens putative peroxisomal antioxidant enzyme mRNA, complete cds.

ACCESSION AF112212
VERSION AF112212.1 GI:6563211

KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Ren, S., Shi, J., Huang, C., Jiang, C., Li, Y., Zhou, J., Yu, Y., Xu, S.,
Wang, Y., Fu, G., Chen, Z. and Han, Z.

TITLE A novel gene expressed in human adrenal gland

REFERENCE
AUTHORS Unpublished
Ren, S., Shi, J., Huang, C., Jiang, C., Li, Y., Zhou, J., Yu, Y., Xu, S.,
Wang, Y., Fu, G., Chen, Z. and Han, Z.

TITLE Direct Submission
JOURNAL Submitted (10-DEC-1998) Chinese National Human Genome Center at
Shanghai, Guo Shoujing Rd. 351, Zhangjiang Hi-tech Park, Pudong,
Shanghai 201203, P. R. China

FEATURES
source Location/Qualifiers
1..835
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
184..672
/note="PLP"

CDS
/codon_start=1
/evidence=not experimental
/product="putative peroxisomal antioxidant enzyme"
/protein_id="AA17200.1"
/db_xref="GI:6563212"

TPCSTKTHLPFGVEQALPAKQVQVACLSVNDFAVTGEWGRAHRAEGKVRLLADPT
GAFGKKTLLILDDSLVSIIFGNRLKRFSMVYQDGIKALNVEPDGTLTCLSLAPNIIIS
QL"

PUBMED 10514471
 REFERENCE 2 (bases 1 to 840)
 AUTHORS Yamashita, H., Avraham, S., London, R., Van Veidhoven, P.P.,
 Subramani, S. and Avraham, H.
 TITLE Direct Submission
 JOURNAL Submitted (31-JAN-1999) Experimental Medicine, Beth Israel
 Deaconess Medical Center, 4 Blackfan Circle, Boston, MA 02115, USA
 FEATURES Location/Qualifiers
 source 1..840
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 CDS 226..714
 /note="PMP20"
 /codon_start=1
 /product="peroxisomal membrane protein 20"
 /protein_id="AAF27531.1"
 /db_xref="GI:6746355"
 /translation="MAPIKVGDAIPAVEVEFEGEPGNKVNLAELFKGKGVLFVGPVAF
 TPQSKTHLPFGVEQAEALKAKGVQVACLVSNDAFVTGEGRAHKAEGKVRLLADPT
 GAFGETDLLDDSLVSIFGNRLKRFMSVQDGVKALNVEPDGTGLTCSLAPNIIS
 QL"

BASE COUNT 166 a 214 c 289 g 171 t
 ORIGIN

alignment_scores:
 Quality: 830.00 Length: 162
 Ratio: 5.123 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-486-167A-2 x AF124993

Align seg 1/1 to: AF124993 from: 1 to: 840

1 MetAlaProLLeLysValGlyAspAlaIleProAlaValGluValPheG1 17
 226 ATGCCCAATCAAGTGGGAGATGCCATCCACAGAGTGGAGGTGTTGA 275
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysLysL 34
 276 AGGGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 325
 34 ySgLyValLeuPheGlyValProGlyValAlaPheThrProGlyCysSerLys 50
 326 AGGTGTGCTGTTGGAGTCTCTGGGGCTTCACCCCTGGATGTTCCAAG 375
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
 376 ACACACTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTCAAGGCCAAGG 425
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 426 AGTCCAGGTGGTGGCTGTGAGTGTATGATGCTTTGTGACTGGCG 475
 84 luTrpGlyArgAlaHisLysAlaGluLysValArgLeuLeuAlaAsp 100
 476 AGTGGGGCCGAGCCACACAGCGGAGGCAAGTTCCGGCTCTCGCTGAT 525
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117
 526 CCACCTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGCT 575
 117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
 576 GGTGTCATCTTTGGGAATCGACGTCCTCAAGAGGTTCCTCATGGGTAC 625
 134 InAspGlyLleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
 626 AGGATGGCATACTGAAGGCCCTGATGTGGAACACAGATGGCACAGCCCTC 675
 151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162

676 ACCTGCAGCTGGCACCACCAATATCATCTCACAGCTC 711

seq_name: gb_pr5:AF242525

seq_documentation_block:

LOCUS AF242525 875 bp mRNA PRI 15-AUG-2000
 DEFINITION Homo sapiens hypothetical protein SBB110 mRNA, complete cds.
 ACCESSION AF242525
 VERSION AF242525.1 GI:9802047

KEYWORDS human.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 875)

AUTHORS Zhang, W., Li, N., Wan, T. and Cao, X.

TITLE Direct Submission

JOURNAL Submitted (08-MAR-2000) Department of Immunology, Second Military

Medical University & Shanghai Brilliance Biotechnology Institute,

800 Xiangyin Rd., Shanghai 200433, P.R. China

FEATURES Location/Qualifiers

1..875

/organism="Homo sapiens"

/db_xref="taxon:9606"

36..680

/codon_start=1

/product="hypothetical protein SBB110"

/protein_id="AAF99605.1"

/db_xref="GI:9802048"

/translation="MGLAGVLCALRRSAGYILVGGAGQSAAAAARCSGEWASGGVR

SFSRAAAMAAPIKVDAIPAVEVEFEGEPGNKVNLAELFKGKGVLFVGPVAFPGCSK

THLPFGVEQAEALKAKGVQVACLVSNDAFVTGEGRAHKAEGKVRLLADPTGAFGKE

TDLLDDSLVSIFGNRLKRFMSVQDGVKALNVEPDGTGLTCSLAPNIISOL"

BASE COUNT 230 a 206 c 272 g 167 t

ORIGIN

alignment_scores:

Quality: 830.00 Length: 162
 Ratio: 5.123 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-486-167A-2 x AF242525

Align seg 1/1 to: AF242525 from: 1 to: 875

1 MetAlaProLLeLysValGlyAspAlaIleProAlaValGluValPheG1 17
 192 ATGCCCAATCAAGTGGGAGATGCCATCCACAGAGTGGAGGTGTTGA 241
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysLysL 34
 242 AGGGAGCCAGGGAACAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 291
 34 ySgLyValLeuPheGlyValProGlyValAlaPheThrProGlyCysSerLys 50
 292 AGGTGTGCTGTTGGAGTCTCTGGGGCTTCACCCCTGGATGTTCCAAG 341
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
 342 ACACACTGCCAGGGTTTGTGGAGCAGGCTGAGGCTCTCAAGGCCAAGG 391
 67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 392 AGTCCAGGTGGTGGCTGTGAGTGTATGATGCTTTGTGACTGGCG 441
 84 luTrpGlyArgAlaHisLysAlaGluLysValArgLeuLeuAlaAsp 100
 442 AGTGGGGCCGAGCCACACAGCGGAGGCAAGTTCCGGCTCTCGCTGAT 491
 101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117

BASE COUNT 141 a 189 c 245 g 150 t
ORIGIN
/note="probable alternative translation start site"
alignment_scores:
Quality: 822.00 Length: 162
Ratio: 5.074 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.765
alignment_block:
US-09-486-167A-2 x AF110736 ..
Align seg 1/1 to: AF110736 from: 1 to: 725

1 MetAlaProIleLeuValGlyAspAlaIleProAlaValGluValPheG1 17
164 ATGCCCCCATCAAGTGGAGATGCCATCCCTGCAGTGGAGGTGTTGA 213
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
214 AGGGAGCCAGGCAACAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 263
34 yGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
264 AGGGTGTGCTGTTGGAGTTCCCGGGGCTTCACGCTGGATGTTCCAG 313
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
314 ACCACCTACCAAGGTTTGTGGAGCAGGCTGAGGCTCTGAAGGCCAAGG 363
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
364 AGTCCAGGTGTGGCTGTGAGTGTAAATGATGCTTTGTGACTGGCG 413
84 lUrPGLyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
414 AGTGGGCCCGAGCCACAGGCAAGGCAAGGTTCCGCTCCTGGCTGAT 463
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
464 CCCACTGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCGTT 513
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
514 AGTGTCCATCTTTGGGAATCGACGCTCAAGAGGTTCTCCATGGTGTAC 563
134 lAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
564 AGGATGCATAGTGAAGCCCTGAATGTGGAACCATGACAGATGGCACAGGCTC 613
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
614 ACCTGCAGTGTGGCAGCCAGCATCATCTCACAGCTC 649

seq_name: gb_pr5:AF197952

seq_documentation_block:
LOCUS AF197952 780 bp mRNA PRI 29-FEB-2000
DEFINITION Homo sapiens thioredoxin peroxidase PMP20 mRNA, complete cds.
ACCESSION AF197952
VERSION AF197952.1 GI:6166492
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 780)
Zhou, Y., Kok, K.H., Chun, A.C., Wong, C.M., Wu, H.W., Lin, M.C.,
Fung, P.C., Kung, H. and Jin, D.Y.
Mouse peroxiredoxin V is a thioredoxin peroxidase that inhibits
p53-induced apoptosis
Biochem. Biophys. Res. Commun. 268 (3), 921-927 (2000)

492 CCACACTGGGGCTTTGGGAAGAGACAGACTTATTACTAGATGATTCGCT 541
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
542 GTGTCCATCTTTGGGAATCGAGCTCAAGAGGTTCTCCATGGTGTAC 591
134 lAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
592 AGGATGGCATAGTGAAGGCCCTGAATGTGGAACCATGTCACAGGCTC 641
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
642 ACCTGCAGCTGTGCCCAATATCATCTCACAGCTC 677

seq_name: gb_pr4:AF110736

seq_documentation_block:
LOCUS AF110736 725 bp mRNA PRI 27-SEP-2000
DEFINITION Cercopithecus aethiops peroxiredoxin 5 (PRDX5) mRNA, complete cds.
ACCESSION AF110736
VERSION AF110736.2 GI:10305337
KEYWORDS
SOURCE African green monkey.
ORGANISM Cercopithecus aethiops
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Cercopithecus.

REFERENCE
AUTHORS Knoop, B., Clippe, A., Bogard, C., Aarslane, K., Wattiez, R.,
Hermans, C., Duconselle, E., Falmagne, P. and Bernard, A.
TITLE Cloning and characterization of AOEB166, a novel mammalian
antioxidant enzyme of the peroxiredoxin family
J. Biol. Chem. 274 (43), 30451-30458 (1999)
JOURNAL 99452929
MEDLINE 10521424
PUBMED
REFERENCE
AUTHORS Knoop, B. and Cherif, H.
TITLE Cloning and characterization of COS-7 AOEB166/PRDX5
unpublished

REFERENCE
AUTHORS Knoop, B., Wattiez, R., Falmagne, P., Hermans, C. and Bernard, A.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-1998) Department of Biology, Catholic University
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
REFERENCE
AUTHORS Knoop, B., Wattiez, R., Falmagne, P., Hermans, C. and Bernard, A.
TITLE Direct Submission
JOURNAL Submitted (26-SEP-2000) Department of Biology, Catholic University
of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
REMARK Sequence update by submitter
COMMENT On Sep 26, 2000 this sequence version replaced gi:10129963.
FEATURES
Location/Qualifiers

1..725
/organism="Cercopithecus aethiops"
/db_xref="taxon:9534"
/cell_line="COS-7"
1..725
/gene="PRDX5"
5..652
/gene="PRDX5"
/note="AOEB166; PRDX5; alkyl hydroperoxide reductase;
thioredoxin peroxidase; mitochondrial, peroxisomal, and
cytosolic"
/codon_start=1
/product="peroxiredoxin 5"
/protein_id="AAG13453.2"
/db_xref="GI:10305338"
/translation="MGIAGVCVLRSSAGYILGGAARQSVAAATAAARRRSGGWSGV
RFSFAAAMAPIKVGDAIPAVEFEGEGNKNVLAELFKRGKGVLFVPGFTPCGS
KTHLPQFVEQALAKGVQVLACSLNDFAVFTGEGWRAHKAEGKRVLLADPTGAEGK
ETDILLDDLSIVSIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPSIISQL"
misc_feature
164
/gene="PRDX5"

source
gene
CDS

```

MEDLINE 20145535
REFERENCE 2 (bases 1 to 780)
AUTHORS Jin,D.-Y.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1999) Institute of Molecular Biology, The
        University of Hong Kong, 8 Sassoon Road, Hong Kong
FEATURES
    source      Location/Qualifiers
                1..780
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="11"
                /map="11q13"
                /note="type V peroxidase"
                /codon_start=1
                /product="thioredoxin peroxidase PMP20"
                /protein_id="AAF04856.1"
                /db_xref="GI:6166493"
                /translation="MGLAGVCLRRSAGYILVGGAGGQSAARRCSEGEWASGGVR
                SFSSRAAAMAPIKVGDAPAVEVEFEGEPGNKNLAEFLKGGKGVLFVPGFTPGCSK
                THLPFGVEQAELKAGVQVACLSDVNDFAVTGEMGRATKAEKGVRLADPTGATGCKE
                TDLLDDSLVSIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNIIISQL"
BASE COUNT 147 a 203 c 278 g 152 t
ORIGIN
alignment_scores:
    Quality: 820.00      Length: 162
    Ratio: 5.093        Gaps: 0
    Percent Similarity: 99.383      Percent Identity: 99.383
alignment_block:
US-09-486-167A-2 x AF197952 ..
Align seg 1/1 to: AF197952 from: 1 to: 780
1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
|||||
266 ATGGCCCAATCAAGTGGGAGATGCCATCCAGCAGTGGAGGTGTTGA 315
|||||
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
316 AGGGGAGCCAGGACACAGGTGAACCTGGCAGAGCTGTTCAAGGGCAGA 365
|||||
34 ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
366 AGGTGTGCTGTTGGAGTCTCTGGGCGCTTCAACCCCTGGATGTTCCAAG 415
|||||
51 ThrHisLeuProGlyPheValGluAlaGluAlaLeuLysAlaLysG1 67
|||||
416 ACACACCTGCCAGGCTTTGGAGCAGCGCTGAGGCTCTGAAGGCCAAGG 465
|||||
67 yValGlnValValAlaLysLeuSerValAsnAspAlaPheValThrGlyG 84
|||||
466 AGTCCAGGTGTGGCTGCTGAGCTGTTAAGATGCGCTTTGTGACTGGCG 515
|||||
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||
516 AGTGGGCGCGACCAACAAAGCGGAAGGATTCGGCTCTCGCTGAT 565
|||||
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117
|||||
566 CCCACTGGGCGCTTTGGAGGAGACAGACTTATCTAGATGATCGCT 615
|||||
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
|||||
616 GGTGTCCATCTTTGGGAATCGAGCTCTCAAGAGGTCTCTCCATGGTGTAC 665
|||||
134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
666 AGGATGCATAGTAGAGCGCTTGAATGTGAACACAGATGGCACAGGCGCTC 715
|||||
151 ThrCysSerLeuAlaProAsnIleSerGlnLeu 162

```

```

|||||
716 ACCTGCAGCCTGGCACCAATATATCATCTCACAGCTC 751
seq_name: gb_pr4:AF110734
seq_documentation_block:
LOCUS AF110734 725 bp mRNA PRI 27-SEP-2000
DEFINITION Papio hamadryas peroxidexin 5 (PRDX5) mRNA, complete cds.
ACCESSION AF110734
VERSION AF110734.2 GI:10305333
KEYWORDS
SOURCE baboon.
ORGANISM
    Papio hamadryas
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae;
    Cercopithecinae; Papio.
REFERENCE
    1 (bases 1 to 725)
    Knoops,B., Clippe,A., Bogard,C., Arsalane,K., Wattiez,R.,
    Hermans,C., Duconseille,E., Falmagne,P. and Bernard,A.
    Cloning and characterization of AOEB166, a novel mammalian
    antioxidant enzyme of the peroxidexin family
    J. Biol. Chem. 274 (43), 30451-30458 (1999)
MEDLINE 99452929
PUBMED 10521424
REFERENCE
    2 (bases 1 to 725)
    Knoops,B. and Cherif,H.
    Cloning and characterization of baboon AOEB166/PRDX5
    Unpublished
    3 (bases 1 to 725)
    Knoops,B., Wattiez,R., Falmagne,P., Hermans,C. and Bernard,A.
    Direct Submission
    Submitted (04-DEC-1998) Department of Biology, Catholic University
    of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
    4 (bases 1 to 725)
    Knoops,B., Wattiez,R., Falmagne,P., Hermans,C. and Bernard,A.
    Direct Submission
    Submitted (26-SEP-2000) Department of Biology, Catholic University
    of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
REMARK
    Sequence update by submitter
COMMENT
    On Sep 26, 2000 this sequence version replaced gi:10129958.
FEATURES
    source
    1..725
    /organism="Papio hamadryas"
    /db_xref="taxon:9557"
    1..725
    /gene="PRDX5"
    5..652
    /gene="PRDX5"
    /note="AOEB166; PRDX5; alkyl hydroperoxide reductase;
    thioredoxin peroxidase; mitochondrial, peroxisomal, and
    cytosolic"
    /codon_start=1
    /product="peroxidexin 5"
    /protein_id="AAG13451.2"
    /db_xref="GI:10305334"
    /translation="MGLAGVCLRRSAGYILGGAAGOSVAATAARRRRSGGWASGGV
    RSFSRAAAMAPIKVGDAPAVEVEFEGEPGNKNLAEFLKGGKGVLFVPGFTPGCS
    KTHLPFGVEQAELKAGVQVACLSDVNDFAVTGEMGRATKAEKGVRLADPTGATGCK
    ETDLLDDSLVSIFGNRLKRFMSVVDGIVKALNVEPDGTGLTCSLAPNIIISQL"
    164
    /gene="PRDX5"
    /note="probable alternative translation start site"
BASE COUNT 141 a 186 c 246 g 152 t
ORIGIN
alignment_scores:
    Quality: 818.00      Length: 162
    Ratio: 5.049        Gaps: 0
    Percent Similarity: 100.000      Percent Identity: 98.148
alignment_block:
US-09-486-167A-2 x AF110734 ..

```

Align seg 1/1 to: AF110734 from: 1 to: 725

```

1  MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
164  ATGGCCCCCGATCAAGTGGAGATGCCATCCCTGCAGTGGAGGTGTTGA 213
17  uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
214  AGGGAGCCGACGGACAAAGGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 263
34  ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
264  AGGGTGTCTGTTGGAGTTCCTGGGGGCTTCACGCCCTGGATGTTCCAAG 313
51  ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
314  ACCCACTACCAGGGTTTGTGGAGCAGCTGAGGCTCTGAAGGCCCAAGG 363
67  yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
364  AGTCAGGTGTGGCTCTCTGAGTGTAAATGATGCCCTTGTGACTGGCG 413
84  LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
414  AGTGGGCGGAGCCACAAAGGTGAAGCAAGGTTCCGGCTCTGCTGCTGAT 463
101  ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
464  CCCACTGGGGCTTTGGGAAGGAGACAGACTTATTACTAGATGATTCATT 513
117  uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
514  AGTGCCATCTTTGGGAATCGACGCTCAAGAGGTTCTCCATGGTGGTAC 563
134  InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
564  AGGATGGCATAGTGAAGGCCCTGAATGTGAACACAGATGGCACAGGCC 613
151  ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
614  ACCTGCAGTCTGGCACCCACCATCATCTCACAGCTC 649

```

seq_name: gb_pat1.A99101

```

seq_documentation_block: 675 bp DNA PAT 20-SEP-2000
LOCUS A99101
DEFINITION Sequence 5 from Patent WO9909054.
ACCESSION A99101
VERSION A99101.1 GI:6782051
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 675)
FALMAGNE, P., WATTIEZ, R., BERNARD, A., HERMANS, C. and KNOOPS, B.
PEROXISOME-ASSOCIATED POLYPEPTIDE, NUCLEOTIDE SEQUENCE ENCODING
SAID POLYPEPTIDE AND THEIR USES IN THE DIAGNOSIS AND/OR THE
TREATMENT OF LUNG INJURIES AND DISEASES, AND OF OXIDATIVE
STRESS-RELATED DISORDERS
JOURNAL Patent: WO 9909054-A 5 25-FEB-1999;
UNIV MONS HAINAUT (BE) ; FALMAGNE PAUL (BE) ; WATTIEZ RUDDY (BE) ;
BERNARD ALFRED (BE) ; HERMANS CEDRIC (BE) ; KNOOPS BERNARD (BE) ;
UNIV LOUVAIN (BE)
FEATURES Location/Qualifiers
source 1..675
/organism="Mus sp."
/db_xref="taxon:10095"
99..588
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB70591.1"
CDS

```

/db_xref="GI:6782052"
 /translation="MAPIKVGDAIPSEVEFEGPKKVLAEFLKGGKVLFGVPGAF
 TPGCKSTHLPGFVEQAGALKAKAQVACLVDNFVIEWGRAHOAEKVRLLADPT
 GAGKATDILLDDLSLVSLEFNRRLKRFSSVIDNGIYKALNVDPDGTGLTCSLAPNLLS
 QL"

BASE COUNT 140 a 174 c 214 g 147 t
 ORIGIN

alignment_scores:
 Quality: 767.00 Length: 162
 Ratio: 4.854 Gaps: 0
 Percent Similarity: 97.531 Percent Identity: 91.358

alignment_block:

US-09-486-167A-2 x A99101 ..

Align seg 1/1 to: A99101 from: 1 to: 675

```

1  MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
99  ATGGCCCCCGATCAAGTGGAGATGCCATCCCTGCAGTGGAGGTATTGA 148
17  uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
149  AGGGCAACCGGAAAGAGTGAACCTGGCAGAGCTGTTCAAGGGCAAGA 198
34  ysGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
199  AAGGTGTTTTTTTGGAGTCCCTGGGGCATTTACACCTGGCTGTTCTTAAG 248
51  ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
249  ACCCACTGGCTGGGTTTGTGGAGCAAGCTGGAGCTCTGAAGGCTAAGGG 298
67  yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
299  AGCGCAGGTGGTGGCTGCTGCTGAGCGTTAATGACGCTCTTTGTGATTGAAG 348
84  LuTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
349  AGTGGGTGAGCCACCCAGGAGGAGGCAAGGTTCCGGCTCCTGGCTGAC 398
101  ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
399  CCCACTGGAGCCTTTGGGAAGGCGACAGACTTATTATTGGATGATCTTT 448
117  uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
449  GGTGCTCTCTTTGGGAATCGTGGCTGAAAGAGTTCTCCATGGTGATAG 498
134  InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
499  ACAAGCCGATAGTAGAGGCACTGAAGCTGAGCCAGCATGGCACAGGCC 548
151  ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
549  ACCTGCAGCTGGCCCCCAACATCCTCTCCCAACTC 584

```

seq_name: gb_rol.AF110733

seq_documentation_block:
 LOCUS AF110733 675 bp mRNA ROD 27-SEP-2000
 DEFINITION Mus musculus peroxiredoxin 5 (Prdx5) mRNA, complete cds.

ACCESSION AF110733

VERSION AF110733.1 GI:10129956

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 675)

AUTHORS Knoop, B., Clippe, A., Bogard, C., Arsalane, K., Wattiez, R.,

Hermans, C., Duconseille, E., Falmagne, P. and Bernard, A.
 Cloning and characterization of AOEB166, a novel mammalian
 antioxidant enzyme of the peroxidoxin family
 J. Biol. Chem. 274 (43), 30451-30458 (1999)
 99452929
 PUBMED 10521424
 REFERENCE 2 (bases 1 to 675)
 AUTHORS Knoops, B., Clippe, A. and Bernard, A.
 TITLE Cloning and characterization of mouse AOEB166/PRDX5
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 675)
 AUTHORS Knoops, B., Wattiez, R., Falmagne, P., Hermans, C. and Bernard, A.
 TITLE Direct Submission
 JOURNAL Submitted (04-DEC-1998) Department of Biology, Catholic University
 of Louvain, Place Croix du Sud, 5, Louvain-la-Neuve 1348, Belgium
 FEATURES
 Location/Qualifiers
 1..675
 /organism="Mus musculus"
 /strain="C3H/HeJ"
 /db_xref="taxon:10090"
 /tissue_type="lung"
 1..675
 /gene="Prdx5"
 99..587
 /gene="Prdx5"
 /note="AOEB166; PRDX5; alkyl hydroperoxide reductase;
 thioredoxin peroxidase; peroxisomal and cytosolic"
 /codon_start=1
 /product="peroxidoxin 5"
 /protein_id="AAG13450.1"
 /db_xref="GI:10129957"
 /translation="MAPIKVGDALPSVEFEGEGKVKVNLAELEFKKKGVLFVPGAF
 TPQSKTHLPFGVQAGALKAKGAQVAVCLSVNDVFIEEMGRAHQAEKVRLLADPT
 GAFKATDLLDDLSVLFGRNRLKRFMSVIDNGIVKALNVEPDGTGLTCLAPNLS
 QL"
 BASE COUNT 140 a 174 c 214 g 147 t
 ORIGIN
 alignment_scores
 Quality: 767.00 Length: 162
 Ratio: 4.854 Gaps: 0
 Percent Similarity: 97.531 Percent Identity: 91.358
 alignment_block
 US-09-486-167A-2 x AF110733 ..
 Align seg 1/1 to: AF110733 from: 1 to: 675
 1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
 |||||
 99 ATGGCCCCGATCAAGTGGGAGATGCCATTCCTCAGTGGAGTATTGA 148
 |||||
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34
 |||||
 149 AGGGGAACCGGAAAGAGAGTGGAGCTGTCAGGAGCTTCAAGGGCAAGA 198
 |||||
 34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
 |||||
 199 AAGGTGTTTGTGGAGTCCCTGGGGATTTACACCTGGCTGTCTTAAG 248
 |||||
 51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
 |||||
 249 ACCCAGCTGCCGTGGTGGAGCAAGCTGGAGCTCTGAAGGCTAAGGG 298
 |||||
 67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
 |||||
 299 ACCGAGGTGGGCTGTCTGAGCGTTAATCAGCTCTTGTGATTGAAG 348
 |||||
 84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaasp 100
 |||||
 349 AGTGGGGTCGAGCCACCAGCAGCAGAGGAGGTTTCGGCTCCTGGCTGAC 398
 |||||

101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuAspSerLe 117
 |||||
 399 CCACCTGGAGCCTTTGGGAGCGGACAGACTTATTATTGGATGATTCTTT 448
 |||||
 117 uValSerIlePheGlyAsnArgLeuLysArgPheSerMetValValG 134
 |||||
 449 GGTGTCTCTTTTGGGAATCGTCGGCTGAAAGAGTCTCCATGGTGATAG 498
 |||||
 134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
 |||||
 499 ACAACGGCATAGTAGGACCACTGAACGCTGGAGCCAGATGCGACAGGCCTC 548
 |||||
 151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
 |||||
 549 ACCTGCAGCTGGCCCCCAACATCTCTCCCAACTC 584
 |||||
 seq_name: gb_rol:AF197951
 seq_documentation_block:
 LOCUS AF197951 688 bp mRNA ROD 29-FEB-2000
 DEFINITION Mus musculus thioredoxin peroxidase PMP20 mRNA, complete cds.
 ACCESSION AF197951
 VERSION AF197951.1 GI:6166490
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 688)
 AUTHORS Zhou, Y., Kok, K.H., Chun, A.C., Wong, C.M., Wu, H.W., Lin, M.C.,
 Fung, P.C., Kung, H. and Jin, D.Y.
 TITLE Mouse peroxidoxin V is a thioredoxin peroxidase that inhibits
 p53-induced apoptosis
 Biochem. Biophys. Res. Commun. 268 (3), 921-927 (2000)
 JOURNAL 20145535
 MEDLINE
 REFERENCE 2 (bases 1 to 688)
 AUTHORS Jin, D.-Y.
 TITLE Direct Submission
 JOURNAL Submitted (23-OCT-1999) Institute of Molecular Biology, The
 University of Hong Kong, 8 Sassoon Road, Hong Kong
 FEATURES
 Location/Qualifiers
 1..688
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 29..661
 /note="type V peroxidoxin"
 /codon_start=1
 /product="thioredoxin peroxidase PMP20"
 /protein_id="AA04855.1"
 /db_xref="GI:6166491"
 /translation="MLQLGLRVLGCKASSVLRASVTLGRAGRKAGWECGGARSSFS
 SAVTMAPIKVGDALPSVEFEGEGKVKVNLAELEFKKKGVLFVPGAFTPCQSKTHLP
 GFVEQAGALKAKGAQVAVCLSVNDVFIEEMGRAHQAEKVRLLADPTGAGKATDLL
 LDLSVLSFGNRLKRFMSVIDNGIVKALNVEPDGTGLTCLAPNLSQL"
 BASE COUNT 138 a 167 c 233 g 150 t
 ORIGIN

alignment_scores
 Quality: 767.00 Length: 162
 Ratio: 4.854 Gaps: 0
 Percent Similarity: 97.531 Percent Identity: 91.358
 alignment_block
 US-09-486-167A-2 x AF197951 ..
 Align seg 1/1 to: AF197951 from: 1 to: 688
 1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
 |||||
 173 ATGGCCCCGATCAAGTGGGAGATGCCATTCCTCAGTGGAGTATTGA 222
 |||||
 17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34
 |||||

```
/db_xref="GI:6103726"
/translation="MVLRFCLVLSIAGSVLRASATWTCVAGRKRGAGWEGGARS
FSSANVTMAPIKVDRIPTSEVEPEKPKVNLAEIPKDKKGVLFVGPVGFPGCSKT
HLPGFVEOAGALKAKAGQVYVACLVDNFATVEMGRHAGKQVGLADPTGAFGKET
DLDDSLSLFNGRRUKRFMSWIDKGVKALINVEPDTGLTCSLAPNILSQL"
193
misc_feature
193
/note="probable alternative translation start site"
242
variation
242
/note="results in glutamic acid to glycine substitution"
380
/replace="g"
variation
380
/note="results in leucine to proline substitution"
428
/replace="c"
variation
428
/note="results in alanine to valine substitution"
/replace="t"
BASE COUNT 198 a 204 c 256 g 179 t
ORIGIN
alignment_scores:
Quality: 767.00 Length: 162
Ratio: 4.824 Gaps: 0
Percent Similarity: 98.148 Percent Identity: 90.741
alignment_block:
US-09-486-167A-2 x AF110732
Align seg 1/1 to: AF110732 from: 1 to: 837
1 MetAlaProIleLysValGlyAspAlaIleProAlaValGluValPheG1 17
193 ATGGCCCGATCAAGGTGGGAGACACCATTCCTCGTCAAGGATTTGA 242
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyL 34
243 AGGGGAACCTGGAAAGAGGTGAAGTGGAGAGCTGTTCAAGGACAAGA 292
34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
293 AAGGTGTTTGTGGAGTCCCTGGGGCAATTACACCTGGCTGTCCAAG 342
51 ThrHisLeuProGlyPheValGluGlnAlaGluLeuLysAlaLysG1 67
343 ACCCATCTGCTGGTGTGGAGCAAGCCGAGCTCTGAAGGCCAAGGG 392
67 yValGlnValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
393 AGCACAAGTGGTGGCTGTCTGAGTGTAAATGATGCCTTCGTGACTGAG 442
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
443 AGTGGGTCGAGCCACAGCAGCAGAGGAGGTTTCAGCTCTCGCTGAC 492
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
493 CCCACTGGAGCTTTTGAAGAGGACAGACATTTACTAGATGATCTTT 542
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValValG 134
543 GGTCTCTCTCTTTGGGAATCGCTGGCTAAAAAGGTTCTCCATGGTGATAG 592
134 lnAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
593 ACAAGGGCGTAGTAAGGCACTGAACGTGGAGCCGGATGGCAGGCCTC 642
151 ThrCysSerLeuAlaProAsnIleIleSerGlnLeu 162
643 ACCTGCAGCCTGGCCCCCAACATCTCTCTCAACAATC 678
seq_name: gb_rol:AF124994
seq_documentation_block:
```

```

LOCUS       AF124994      895 bp      mRNA                      25-JAN-2000
DEFINITION  Mus musculus peroxisomal membrane protein 20 mRNA, complete cds.
ACCESSION   AF124994
VERSION     AF124994.1  GI:6746356
KEYWORDS    house mouse.
SOURCE      Mus musculus
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 895)
AUTHORS     Yamashita,H., Avraham,S., Jiang,S., London,R., Van Veldhoven,P.P.,
            Subramani,S., Rogers,R.A. and Avraham,H.
TITLE       Characterization of human and murine PMP20 peroxisomal proteins
            that exhibit antioxidant activity in vitro
JOURNAL     J. Biol. Chem. 274 (42), 29897-29904 (1999)
MEDLINE     99445545
PUBMED      10514471
REFERENCE   2 (bases 1 to 895)
AUTHORS     Yamashita,H., Avraham,S., London,R., Van Veldhoven,P.P.,
            Subramani,S. and Avraham,H.
TITLE       Direct Submission
JOURNAL     Submitted (31-JAN-1999) Experimental Medicine, Beth Israel
            Deaconess Medical Center, 4 Blackfan Circle, Boston, MA 02115, USA
FEATURES    Location/Qualifiers
             1..895
             /organism="Mus musculus"
             /db_xref="taxon:10090"
             262..750
             /note="PMP20"
             /codon_start=1
             /product="peroxisomal membrane protein 20"
             /protein_id="AAF27532.1"
             /db_xref="GI:6746357"
             /translation="MAPIKVGDALPSVEFEPEGPKKVNLAELFKGKGVLFVPGAF
             TPQSKTHLPFGVQAGALKAGVAVCLSVNDVFVIEWGRAHQAEGKVRLLADPT
             GAFGKATDLLLDSLSLFGNRLKRFMSWIDNGIVKALNVEPDGTLTCSLAPNILS
             QL"
BASE COUNT  181 a 232 c 284 g 198 t
ORIGIN
alignment_scores:
  Quality: 767.00      Length: 162
  Ratio: 4.854        Gaps: 0
Percent Similarity: 97.531 Percent Identity: 91.358
alignment_block:
US-09-486-167A-2 x AF124994  ..
Align seg 1/1 to: AF124994 from: 1 to: 895
1 MetAlaProIleIysValGlyAspAlaIleProAlaValGluValPheG1 17
|||||
262 ATGGCCCCGATCAAGTGGGAGATCCATCCCTCAGTGGAGGTATTGA 311
17 uGlyGluProGlyAsnLysValAsnLeuAlaGluLeuPheLysGlyLysL 34
|||||
312 AGGGGAACCGGAAGAGTGAAGTGGCAGAGCTGTTCAAGGGCAAGA 361
34 ySGlyValLeuPheGlyValProGlyAlaPheThrProGlyCysSerLys 50
|||||
362 AAGGTGTTTGTGAGTCCCTGGGGCATTTACACCTGGCTGTTCTAAG 411
51 ThrHisLeuProGlyPheValGluGlnAlaGluAlaLeuLysAlaLysG1 67
|||||
412 ACCCACTCCCTGGGTTTGTGGAGCAAGCTGTGAAGCCCAAGGG 461
67 yValGlnValValAlaCysLeuSerValAsnAspAlaPheValThrGlyG 84
|:::|
462 ACCGAGGTGGTGGCTGCTCTGAGGCTTAATGACGCTTTGTGATTGAAG 511
84 luTrpGlyArgAlaHisLysAlaGluGlyLysValArgLeuLeuAlaAsp 100
|||||

```

```

512 AGTGGGGTCGAGCCACCAGGAGGCAAGGTTTCGGCTCCTGGCTGAC 561
101 ProThrGlyAlaPheGlyLysGluThrAspLeuLeuLeuAspSerLe 117
|||||
562 CCCACTGGAGCCTTTGGGAAGCGCAGACTTATTATTGGATGATCTTT 611
117 uValSerIlePheGlyAsnArgArgLeuLysArgPheSerMetValG1 134
|||||
612 GGTGTCCTCTTTGGGAATCGTCGGCTGAAAAGGTTCTCCATGGTGATAG 661
134 InAspGlyIleValLysAlaLeuAsnValGluProAspGlyThrGlyLeu 150
|||||
662 ACAACGGCATAGTAGAAGCACTGAACGTGGAGCCAGATGGCACAGGCCTC 711
151 ThrCysSerLeuAlaProAsnIleLeuSerGlnLeu 162
|||||
712 ACCTGCGAGCTGGCCCCCAACATCTCTCTCAATC 747

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 22, 2001, 10:44:35 ; Search time 58.22 Seconds
(without alignments)
2617.579 Million cell updates/sec

Title: US-09-486-167a-1

Perfect score: 805

Sequence: 1 gccaggagcggagtggaag.....ttgtggtttgcggaagaaaaa 805

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

- 1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	798	99.1	993	4	US-08-959-004-2
2	49.4	6.1	7218	1	US-08-232-463-14
3	39.4	4.9	4616	1	US-08-340-203A-1
4	39.4	4.9	4616	2	US-08-452-567-1
5	39.4	4.9	4616	2	US-08-452-567-1
6	39.4	4.9	4616	3	US-09-085-407-1
7	37.2	4.6	4112	1	US-08-340-203A-2
8	37.2	4.6	4112	2	US-08-452-567-2
9	37.2	4.6	4112	3	US-09-085-407-2
10	37.2	4.6	4112	3	US-09-085-407-2
11	36.2	4.5	68750	3	US-09-335-409-1
12	36	4.5	458	3	US-09-141-000-4
13	34.8	4.3	418	3	US-09-141-000-6
14	34	4.2	4252	2	US-08-475-844-4
15	34	4.2	4252	5	PCT-US95-08429-4
16	33.6	4.2	2219	3	US-08-510-646B-17
17	33.6	4.2	6453	1	US-08-306-691B-14
18	33.6	4.2	6453	3	US-09-209-668-10
19	33.6	4.2	6453	3	US-09-356-952-8
20	33.4	4.1	1611	2	US-08-650-598-1
21	33.4	4.1	35060	3	US-08-814-095-7
22	33.2	4.1	2329	1	US-08-457-797A-8
23	33.2	4.1	2329	1	US-08-812-025-8
24	33.2	4.1	2721	6	5215881-2
25	33.2	4.1	8438	1	US-07-945-283-1
26	32.8	4.1	1866	4	US-09-318-448-33
27	32.8	4.1	2384	1	US-07-814-964-10

28	32.8	4.1	2384	1	US-08-258-442-10	Sequence 10, Appl
29	32.8	4.1	2384	1	US-08-328-809-5	Sequence 5, Appl
30	32.8	4.1	2384	5	PCT-US92-11107-10	Sequence 10, Appl
31	32.6	4.0	2031	2	US-08-933-750C-53	Sequence 53, Appl
32	32.6	4.0	2031	3	US-09-234-613-53	Sequence 53, Appl
33	32.4	4.0	12001	1	US-08-458-568A-11	Sequence 11, Appl
34	32.2	4.0	1528	4	US-09-362-473-13	Sequence 13, Appl
35	32.2	4.0	2580	3	US-09-050-863-2	Sequence 2, Appl
36	32.2	4.0	5452	2	US-09-130-114-1	Sequence 1, Appl
37	32.2	4.0	9600	4	US-08-910-647-1	Sequence 1, Appl
38	32.2	4.0	10596	1	US-07-884-811-15	Sequence 15, Appl
39	32.2	4.0	10596	1	US-07-885-971-15	Sequence 15, Appl
40	32.2	4.0	10596	1	US-08-087-783A-15	Sequence 15, Appl
41	32.2	4.0	10596	1	US-08-194-088B-15	Sequence 15, Appl
42	32.2	4.0	10596	2	US-08-194-087-15	Sequence 15, Appl
43	32.2	4.0	10596	5	PCT-US93-04648-15	Sequence 15, Appl
44	32	4.0	320	4	US-09-165-264-11	Sequence 11, Appl
45	31.6	3.9	319	4	US-09-165-264-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-08-959-004-2
; Sequence 2, Application US/08959004
; Patent No. 6197543

GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Yue, Henry

; APPLICANT: Corley, Neil C.

; APPLICANT: Lal, Preeti

; APPLICANT: Shah, Purvi

; APPLICANT: Kaser, Matthew

; TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/959,004

FILING DATE: Herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0414 US

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 993 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: BRAITUT01

CLONE: 743725


```

; Patent No. 5756668
; GENERAL INFORMATION:
; APPLICANT: Baylin, Stephen B.
; APPLICANT: Wales, Michele M.
; TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,203A
; FILING DATE: 15-NOV-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: 07265/039001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4616 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: HIC-1 polynucleotide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4616
; US-08-340-203A-1

```

```

Query Match      4.9%; Score 39.4; DB 1; Length 4616;
Best Local Similarity 48.4%; Pred. No. 0.091;
Matches 139; Conservative 0; Mismatches 146; Indels 2; Gaps 1;

QY      1  gccaggagcgagtggaagtggcggtggcggtatgggactagctggcggtgagcgc 60
      || || || || || || || || || || || || || || || || || || || || ||
Db      1575  GCGTGGCGCGCCGCGAGGCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1516

QY      61  ctgagacgtcagcggtctatactcgtgtggcggtggcggtggcggtggcggtggcggt 120
      || || || || || || || || || || || || || || || || || || || || ||
Db      1515  CGCCGCCGCCGCGAGGTGGCAGTACTTGCCTGGCGCTTGGCGCTTGGCAGCG 1456

QY      121  gcagcaagcaggtgcagtggaaggagagtg--ggcgctggcggtggcggtggcggtggcggt 178
      || || || || || || || || || || || || || || || || || || || || ||
Db      1455  CCACAGGTTCGGGATCTGCAGGTAGCTGGCGCGCGCGCGCGCGCGCGCGCGCG 1396

QY      179  gagcgctgcagcattggcccaatcaagggtgggagatgccatcccgagtgaggtgt 238
      || || || || || || || || || || || || || || || || || || || || ||
Db      1395  CAGCCCCCGGGCCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1336

QY      239  ttgaaggaggagccaggggaacaggtgaaacctggcagagctgttcaag 285
      || || || || || || || || || || || || || || || || || || || || ||
Db      1335  TGTAGATGAAGTCCAGCAGCAGGCGGACACGCGCGCGCGCGCGCGCGCGCGCGCGCG 1289

```

```

RESULT      4
US-08-452-567-1/c
; Sequence 1, Application US/08452567
; Patent No. 5846712

```

```

; GENERAL INFORMATION:
; APPLICANT: Baylin, Stephen B.
; APPLICANT: Wales, Michele M.
; TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,567
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/039003 (Div PD3664)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4616 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: HIC-1 polynucleotide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4616
; US-08-452-567-1

```

```

Query Match      4.9%; Score 39.4; DB 2; Length 4616;
Best Local Similarity 48.4%; Pred. No. 0.091;
Matches 139; Conservative 0; Mismatches 146; Indels 2; Gaps 1;

QY      1  gccaggagcgagtggaagtggcggtggcggtatgggactagctggcggtgagcgc 60
      || || || || || || || || || || || || || || || || || || || || ||
Db      1575  GCGTGGCGCGCCGCGAGGCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1516

QY      61  ctgagacgtcagcggtctatactcgtgtggcggtggcggtggcggtggcggtggcggt 120
      || || || || || || || || || || || || || || || || || || || || ||
Db      1515  CGCCGCCGCCGCGAGGTGGCAGTACTTGCCTGGCGCTTGGCGCTTGGCAGCG 1456

QY      121  gcagcaagcaggtgcagtggaaggagagtg--ggcgctggcggtggcggtggcggtggcggt 178
      || || || || || || || || || || || || || || || || || || || || ||
Db      1455  CCACAGGTTCGGGATCTGCAGGTAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1396

QY      179  gagcgctgcagcattggcccaatcaagggtgggagatgccatcccgagtgaggtgt 238
      || || || || || || || || || || || || || || || || || || || || ||
Db      1395  CAGCCCCCGGGCCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1336

QY      239  ttgaaggaggagccaggggaacaggtgaaacctggcagagctgttcaag 285
      || || || || || || || || || || || || || || || || || || || || ||
Db      1335  TGTAGATGAAGTCCAGCAGCAGGCGGACACGCGCGCGCGCGCGCGCGCGCGCGCGCG 1289

```

```

RESULT      5
US-08-452-427-1/c
; Sequence 1, Application US/08452427
; Patent No. 5922590
; GENERAL INFORMATION:

```

us-09-486-167a-1.rni

Mon Oct 22 14:46:53 2001

APPLICANT: Baylin, Stephen B.
APPLICANT: Wales, Michele M.
TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: USA
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,427
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/340,203
FILING DATE: 15-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: 07265/039001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4616 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: HIC-1 polynucleotide
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4616

US-08-452-427-1

Query Match 4.9%; Score 39.4; DB 2; Length 4616;
Best Local Similarity 48.4%; Pred. No. 0.091;
Matches 139; Conservative 0; Mismatches 146; Indels 2; Gaps 1;

1 gccaggagcgagtggaagtggcgtggggcggtatggactagctggcgtgtgcgcc 60
Db 1575 GCGTGGCGCGCCGACGCCCGCGCGACCATAGGCGCGTAGCGCGCGCGCGCG 1516
QY 61 ctgagacgtcagcggtatatactcgtcgtggggcggtcagctcagctcgcagcg 120
Db 1515 CG 1456
QY 121 gcagcaagcggtcagtggaaggagagtg--ggcgtcgtggcggttcgcagtttcagca 178
Db 1455 CCACGAGTGGGATCTGAGGTAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1396
QY 179 gacgcgtcgcagcctggcccaatcaaggctgggagatgccatccagcagctgttcaag 238
Db 1395 CAGCCCCGGGCG 1336
QY 239 ttgaagggggagcaggaagtggaacctggcagagctgttcaag 285
Db 1335 TGTATGAAATCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1289

RESULT 6
US-09-085-407-1/c
; Sequence 1, Application US/09085407

Patent No. 6103877

GENERAL INFORMATION:
APPLICANT: Baylin, Stephen B.
APPLICANT: Wales, Michele M.
TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: USA
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/085,407
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/340,203
FILING DATE: 15-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Ph.D., Lisa A.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: 07265/039001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4616 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: HIC-1 polynucleotide
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4616

US-09-085-407-1

Query Match 4.9%; Score 39.4; DB 3; Length 4616;
Best Local Similarity 48.4%; Pred. No. 0.091;
Matches 139; Conservative 0; Mismatches 146; Indels 2; Gaps 1;

1 gccaggagcgagtggaagtggcgtggggcggtatggactagctggcgtgtgcgcc 60
Db 1575 GCGTGGCGCGCCGACGCCCGCGCGACCATAGGCGCGTAGCGCGCGCGCGCG 1516
QY 61 ctgagacgtcagcggtatatactcgtcgtggggcggtcagctcagctcgcagcg 120
Db 1515 CG 1456
QY 121 gcagcaagcggtcagtggaaggagagtg--ggcgtcgtggcggttcgcagtttcagca 178
Db 1455 CCACGAGTGGGATCTGAGGTAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1396
QY 179 gacgcgtcgcagcctggcccaatcaaggctgggagatgccatccagcagctgttcaag 238
Db 1395 CAGCCCCGGGCG 1336
QY 239 ttgaagggggagcaggaagtggaacctggcagagctgttcaag 285
Db 1335 TGTATGAAATCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1289

RESULT 7

```

Query Match      4.6%; Score 37.2; DB 1; Length 4112;
Best Local Similarity 50.4%; Pred. No. 0.36;
Matches 117; Conservative 0; Mismatches 113; Indels 2; Gaps 1;

QY 56 gcgcctgagacgtcagcggctatatactcgtggtggcgcggtcagctctgcgg 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1520 GCCGCGCGCGCGCGCGCGCGAGTGGCAGTACTTGGCGTGGCGTTCGAGCGGTTTCTGCA 1461
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 116 cagcgcgacagcagcgtgcagtgaaagagtg--ggcgtctggcggggtccgcagttt 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1460 CAGCGCCACGAGGTCGGGGATCTCAGAGTACGTGGGGGGCGGCCAGCACGGCGCCAGGCT 1401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 174 cagcagagccgctgcagccatgccccaaatcaaggtyggagatgccatccagcagtgga 233
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1400 CGGCTCAGCCCCCGGGGCCAGCGCGCGCGCGCGCGCTCTGGCGCGTTCAGCCAGGCG 1341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 234 ggtgtttgaaggaggcaggcagggaacnagtgtaacctggcgagagctgttcaag 285
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1340 GCCGCTGTAGTAAAGTCCAGCACCGAGCGGGAACAGCGCGGGCTCACCATG 1289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT      8
US-08-452-567-2/c
; Sequence 2, Application US/08452567
; Patent No. 5846712
; GENERAL INFORMATION:
; APPLICANT: Baylin, Stephen B.

```

```

Query Match          4.6%; Score 37.2; DB 2; Length 4112;
Best Local Similarity 50.4%; Pred. No. 0.36;
Matches 117; Conservative 0; Mismatches 113; Indels 2; Gaps 1;

QY 56 ggcgcctgagcgtcagcggctatatactctgctggggccggcgctcagttcg 115
Db 1520 GCCGCCGCCGCCGCCGCCAGTACTTGCCTGCGCTTGAGCGGTTCTTGCA 1461

QY 116 caqcggaagcagcgtgtcagtgaaaggagtg--ggcgctctggcggtccgcagttt 173
Db 1460 CAGCGCCACAGAGGTCGGGGATCTGCAGGTAGCTGGCGGGGGCCAGCAGCGGCGCCAGGCT 1401

QY 174 cagcagagccgctcagcccatgcccccaatcaaggctggagatgccatccacagtgga 233
Db 1400 CGGCTCAGCCCCCGGGGCCACGGCCGGCCGCTCTGGCGCGTCAGCCAGGCG 1341

QY 234 ggtgtttgaaggaggagccagggaacaaagtgaacctggcagagctgttcaag 285
Db 1340 GCCGCTGTAGATGAAGTCCAGCACCCAGCGGGAACAGCGCGGGCTCAGCATG 1289

RESULT 9
SEQUENCE 2, Application US/0845247
PATENT 5922590
GENERAL INFORMATION:
APPLICANT: Baylin, Stephen B.
APPLICANT: Wales, Michele M.
TITLE OF INVENTION: NOVEL TUMOR SUPPRESSOR GENE, HIC-1
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.

```

STATE: California
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: 09/085,407
APPLICATION NUMBER: US/09/085,407
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/340,203
FILING DATE: 15-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Ph.D., Lisa A.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: 07265/039001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4112 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: HIC-1 coding polynucleotide
FEATURE:
NAME/KEY: CDS
LOCATION: 1086..2726
S-09-085-407-2

	Query Match	4.6%;	Score 37.2;	DB 3;	Length 4112;
	Best Local Similarity	50.4%;	Pred. No. 0.36;	Indels	Gaps
	Matches 117;	Conservative	0;	Mismatches 113;	
QY	56	gcgccctgagacgctcagcggggtatatactcgtgtgggcccggctcagcttcgcgg	115		
DB	1520	GC CGCGCGCGCCGCCCGCAGGTGGCAGTACTGCCGCTGGCCCTTGAGCGCTTCTTGC	1461		
QY	116	cagcggcagcaagacggtgcagtgaaaggagagt--ggcgtctggcgggtctccgcagttt	173		
DB	1460	CAGCGCCACGAGGTCGGGGGATCTCAGTGTACTGGCGCGCCACGACGCGCGCCAGGCT	1401		
QY	174	cagcagagccgctgcagcccatggcccaatacaagtggtggagatgccatcccgacgagtgg	233		
DB	1400	CGGCTCAGCCCGCGGGGCCACGGCGCGGCCGACGCCCTCTGGCGCGCTCAGCCAGGCG	1341		
QY	234	gggttttgaggggagccaggggaacaagtgtaacctggcagagctgtttcaag	285		
DB	1240	CGCGGCTCTACATGAAGTTCACACACGCGCGGAACACGCGCGGGCTACGACATG	1289		

```

RESULT 11
US-09-335-409-1
; Sequence 1, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17

```

Query Match	4.5%	Score 36:	DB 3;	Length 458;
Best Local Similarity	8.1%	pred. No. 0.32;		
Matches 36;	Conservative 152;	Mismatches 253;	Indels 5;	Gaps 1;
QY	56	gcgccttgagagcgtcagcggtgcgtatatactgctggtgggcccgcgcgtcagctgcgg	115	
DB	2	DSVELCUPESFLHYEBELLCRMSNDRHIDSSCSFIFKTEPSSPASITDSVNHHSPEG	61	
QY	116	cagcgcgagcaagcagctgcagtgagagagagtggcgtctcgtgcgggctccgcagttca	175	
DB	62	SDASSSYSTWNGHQLDSPPIYPGAPILGSGPVRKLYDDCSSTIVEDPQTKCEYMLN	121	
QY	176	gcagagcgcgtgcagccatgcgcgcacatcaagtggtggagatgccatccccagcagtg	232	
DB	122	SMPRKCLVCGDILASGHVYASCEACAKFFKRTTQGNIEYSCPATNECETTKRRKRSQ	181	
QY	233	--agtggtttgaaggcgaccaggaacaaagtggaacctggcagagctgttcaaggcca	290	
DB	182	ACRFMKCLKVGMLKEGVRUDRVRGGQRXYKRRIDAENSPLYLNPLQVPAKKPKYKIVSHL	241	
QY	291	gaagggtgtcgtgtttgttgagttccctggggccctcacccctggtgttcccaagacac	350	
DB	242	LVAEPEKIYAMPDPTVPDSDIKALTTCCLADRELWIIIGWAKHIGPFTSLIADOMSL	301	

us-09-486-167a-1.rni

Mon Oct 22 14:46:53 2001

APPLICANT: Neiman, Paul E.
APPLICANT: Klenova, Elena M.
APPLICANT: Goodwin, Graham H.
APPLICANT: Filippova, Galina N.
APPLICANT: Collins, Steven J.
TITLE OF INVENTION: CTCF
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: One Market Plaza
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/475,844
FILING DATE: 08/261,680
CLASSIFICATION: 536
FILING DATE: 17-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14538A-11-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4252 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
ORIGINAL SOURCE: Gallus domesticus
FEATURE:
NAME/KEY: CDS
LOCATION: 679..2865
US-08-475-844-4

Query Match 4.2%; Score 34; DB 2; Length 4252;
Best Local Similarity 45.5%; Pred. No. 2.9;
Matches 121; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 5 ggagcgagtggaagtggcgtggggggtatgggactagctggcggtggtgacctga 64
Db 491 GCGCGCGGGCGGCGGAGTGCCTGCGGAGCGCGGCTGCCCGGACCGCGGGCAGGA 432

QY 65 gaogctcagcgggtatatactgctgggtggcgccggtcagtcgctggcagcgagc 124
Db 431 GCGGGAGGGAGGGGGGAGGAGCGCGCGCGGAGCGCGCGCGCGCGCGCGCGG 372

QY 125 caagcagcgtcagtgaaagagagtggtggcggtggcggtggcggtggcggtggcggt 184
Db 371 GCTGGGAGGGAGGGGGGAGGAGCGCGCGCGGAGCGCGCGCGCGCGCGCGCGG 312

QY 185 ctgacgcatggcccccaatcaagtggtggagatgccatcccgagcagtggtgttgaag 244
Db 311 CTGGGGCGGGCGGCTCACTTCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 226

QY 245 gggagccaggggaacaagtgaaacctg 270
Db 251 GCGGCGCGCGCGCTGAGGGGAGGCTG 226

Search completed: October 22, 2001, 12:13:46
Job time: 5351 sec

RESULT 15
PCT-US95-08429-4/c
Sequence 4, Application PC/TUS9508429
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: CTCF
NUMBER OF SEQUENCES: 21
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08429
FILING DATE: 15-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/261,680
FILING DATE: 17-JUN-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14538A-11-IPC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4252 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
ORIGINAL SOURCE: Gallus domesticus
FEATURE:
NAME/KEY: CDS
LOCATION: 679..2865
PCT-US95-08429-4

Query Match 4.2%; Score 34; DB 5; Length 4252;
Best Local Similarity 45.5%; Pred. No. 2.9;
Matches 121; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 5 ggagcgagtggaagtggcgtggggggtatgggactagctggcggtggtgacctga 64
Db 491 GCGCGCGGGCGGCGGAGTGCCTGCGGAGCGCGGCTGCCCGGACCGCGGGCAGGA 432

QY 65 gaogctcagcgggtatatactgctgggtggcgccggtcagtcgctggcagcgagc 124
Db 431 GCGGGAGGGAGGGGGGAGGAGCGCGCGCGGAGCGCGCGCGCGCGCGCGCGG 372

QY 125 caagcagcgtcagtgaaagagagtggtggcggtggcggtggcggtggcggtggcggt 184
Db 371 GCTGGGAGGGAGGGGGGAGGAGCGCGCGGAGCGCGCGCGCGCGCGCGCGG 312

QY 185 ctgacgcatggcccccaatcaagtggtggagatgccatcccgagcagtggtgttgaag 244
Db 311 CTGGGGCGGGCGGCTCACTTCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 226

QY 245 gggagccaggggaacaagtgaaacctg 270
Db 251 GCGGCGCGCGCGCTGAGGGGAGGCTG 226

Mon Oct 22 14:46:53 2001

us-09-486-167a-1.rni

Page 9

